



DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/48, C12Q 1/70, C07K 14/15, A61K 31/70	A1	(11) Numéro de publication internationale: WO 99/67395 (43) Date de publication internationale: 29 décembre 1999 (29.12.99)
(21) Numéro de la demande internationale: PCT/FR99/01513 (22) Date de dépôt international: 23 juin 1999 (23.06.99) (30) Données relatives à la priorité: 98/07920 23 juin 1998 (23.06.98) FR (71) Déposant (pour tous les Etats désignés sauf US): INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE-INSERM [FR/FR]; 101, rue de Tolbiac, F-75654 Paris Cedex 13 (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): ALLIEL, Patrick, M. [FR/FR]; 4, rue Lazare Carnot, F-92140 Clamart (FR). PERIN, Jean-Pierre [FR/FR]; 182, rue d'Aulnay, F-92350 Le Plessis-Robinson (FR). RIEGER, François [FR/FR]; 38 bis, boulevard de la République, F-92100 Boulogne (FR). (74) Mandataire: CABINET ORES; 6, avenue de Messine, F-75008 Paris (FR).		(81) Etats désignés: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Publiée <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues.</i>

(54) Title: NUCLEIC SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUMAN ENDOGENOUS RETROVIRAL MOTIFS, AND THEIR USES

(54) Titre: FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX ENDOGENES HUMAINS, ET LEURS APPLICATIONS

(57) Abstract

The invention concerns a novel nucleic sequence and deduced protein sequence family with whole or partial human endogenous retroviral motifs. The invention also concerns the detection and/or the use of said nucleic sequences and said corresponding protein sequences or fragments of said sequences, for diagnostic, prophylactic and therapeutic uses, in particular for neuropathological conditions with autoimmune constituent such as multiple sclerosis. Said purified nucleic acid sequences comprise all or part of a sequence coding for a human endogenous retroviral sequence having at least *env*-type retroviral motifs, corresponding to the sequence SEQ ID NO:1 or to a sequence having a homology level with said sequence SEQ ID NO:1 not less than 80 % on more than 190 nucleotides or not less than 70 % on more than 600 nucleotides for *env*-type domains. The invention further concerns the use of the flanking or adjacent sequences of said sequences and controlled by the latter, as diagnostic reagents.

(57) Abrégé

Nouvelle famille de séquences nucléiques et de séquences protéiques déduites, qui présentent des motifs rétroviraux endogènes humains complets ou partiels. Détection et/ou utilisation desdites séquences nucléiques et desdites séquences protéiques correspondantes ou de fragments de ces séquences, dans le cadre d'applications diagnostiques, prophylactiques et thérapeutiques, en particulier pour des neuropathologies à composante autoimmune comme la sclérose en plaques. Lesdites séquences d'acide nucléique purifiées comprennent tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80 % sur plus de 190 nucléotides ou supérieur ou égal à 70 % sur plus de 600 nucléotides pour les domaines de type *env*. Utilisation des séquences flanquantes ou adjacentes desdites séquences et contrôlées par ces dernières, comme réactifs de diagnostic.

09/719,554

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX ENDOGENES HUMAINS, ET LEURS APPLICATIONS

5 La présente invention est relative à une nouvelle famille de séquences nucléiques et de séquences protéiques déduites, qui présentent des motifs rétroviraux endogènes humains complets ou partiels, ainsi que des séquences flanquantes ou adjacentes desdites séquences, et contrôlées par ces dernières : modification de l'expression ou altération de la structure (polyadénylation, épissage alternatif...) desdites séquences flanquantes.

L'invention est également relative à la détection et/ou à l'utilisation desdites séquences nucléiques et desdites séquences protéiques correspondantes, dans le cadre d'applications diagnostiques, prophylactiques et thérapeutiques, en particulier pour des neuropathologies à composante autoimmune comme la sclérose en plaques.

15 L'invention concerne aussi l'obtention de sondes nucléiques double brins et simple brin anti-sens, de ribozymes, aptes à moduler la réplication virale (T.R. Cech, *Science*, 1987, **236**, 1532-1539 ; R.H. Symons, *Trends Biochem. Sci.*, 1989, **14**, 445-450) des molécules recombinantes correspondantes, et des anticorps associés.

Les rétrovirus sont des virus qui se répliquent uniquement en utilisant la voie inverse du traitement classique de l'information génétique. Ce processus, nommé transcription inverse, est médié par une ADN polymérase ARN dépendante ou transcriptase reverse, codée par le gène *pol*. Les rétrovirus codent aussi au minimum pour deux gènes additionnels. Le gène *gag* code pour les protéines du squelette, matrice, nucléocapside et capsid. Le gène *env* code pour les glycoprotéines d'enveloppe.

20 La transcription rétrovirale est régulée par des régions promotrices ou "enhancers", situées dans des régions hautement répétées ou LTR (*Long Terminal Repeat*) et qui sont présentes aux deux extrémités du génome rétroviral.

Lors de l'infection d'une cellule, la polymérase fait une copie ADN du génome ARN ; cette copie peut alors s'intégrer dans le génome humain. Les rétrovirus ne tuent pas les cellules qu'ils infectent, mais au contraire améliorent souvent leur rapidité de croissance. Les rétrovirus peuvent infecter des cellules germinales ou

30

des embryons à un stade précoce ; ils peuvent dans ces conditions, intégrer la lignée germinale et être transmis par transmission mendélienne verticale. ce qui constitue la relation la plus étroite entre un hôte et son parasite. Ces virus endogènes peuvent dégénérer au cours des générations de l'organisme hôte et perdre leurs propriétés initiales. Cependant certains d'entre eux peuvent conserver tout ou partie de leurs propriétés ou des propriétés des motifs les composant, ou encore acquérir de nouvelles propriétés fonctionnelles présentant un avantage pour l'organisme hôte, ce qui expliquerait la préservation de leur séquence.

L'existence de motifs endogènes présentant de longs cadres de lecture ouverts et/ou soumis à une forte pression de sélection peut donc être indicatrice d'une fonction biologique préservée ou acquise, qui peut correspondre à un bénéfice pour l'organisme hôte. Ces séquences rétrovirales peuvent aussi subir, au cours des générations, des modifications discrètes qui vont être à même de réveiller certaines de leurs potentialités et engendrer ou favoriser des processus pathologiques. Il est apparu récemment nécessaire de faire le bilan et d'identifier ces séquences afin de pouvoir évaluer leur impact fonctionnel.

Les séquences rétrovirales endogènes humaines ou HERVs représentent une part importante du génome humain. Ces régions rétrovirales se présentent sous plusieurs formes :

- des structures rétrovirales endogènes complètes associant des motifs *gag*, *pol* et *env*, flanqués de séquences nucléiques répétées, qui montrent une analogie significative avec la structure LTR-*gag-pol-env*-LTR des rétrovirus infectieux,
- des séquences rétrovirales tronquées ; par exemple, les rétrotransposons sont privés de leur domaine *env* et les rétrotransposons ne possèdent pas les régions *env* et LTR.

Jusqu'à présent l'étude de ces régions du génome a été négligée chez l'Homme pour deux raisons essentielles :

- l'existence d'insertions/délétions qui peuvent décaler le cadre de lecture et de mutations qui modifient la séquence. Ces modifications entraînent des altérations de la structure et par conséquent de la fonction biologique de ces motifs.

- l'absence d'associations avérées avec des pathologies humaines.

La connaissance, récente de fragments significativement représentatifs du génome humain et une orientation des recherches vers une étude structure/fonction des motifs rétroviraux endogènes, ont permis de préciser l'intérêt de ces régions. L'implication de séquences endogènes tronquées ou complètes dans des pathologies chez l'animal est documentée ; par exemple leur association avec des processus tumoraux a été clairement mise en évidence (S.K. Chattopadhyay et coll., 1982, *Nature*, **295**, 25-31). Une recherche visant à préciser l'association ou l'influence des HERVs dans des pathologies humaines se justifie donc aujourd'hui.

10 Une classification des éléments HERV a été proposée (Tönjes R.R. et al., *AIDS & Hum. Retrovirol.*, 1996, **13**, S261-S267; A.M. Krieg et al., *FASEB J.*, 1992, **6**, 2537-2544). Elle est basée sur une homologie de ces séquences avec des rétrovirus isolés chez les animaux, à l'aide de sondes rétrovirales hétérologues. En effet, en général, les HERVs présentent relativement peu d'homologie avec des rétrovirus infectieux humains connus.

Les familles de classe I présentent une homologie de séquence avec les rétrovirus de mammifères de type C ; on peut citer notamment la superfamille ERI, proche du virus MuLV (*murine leukemia virus*) et du virus BaEV (*baboon endogenous virus*).

20 Les familles de classe II présentent une homologie de séquence avec les rétrovirus de mammifères de type B tel que le MMTV (*mouse mammary tumour virus*) ou les rétrovirus de type D tel que le SRV (*squirrel monkey retrovirus*).

D'autres familles ont également été décrites ; parmi celles-ci, on peut citer des HERVs qui présentent, de manière exceptionnelle, une homologie partielle avec HTLV-1 (RTVL-H) ou des virus de primates ; HRES-1, par exemple, présente une homologie de séquence avec des HTLVs.

Les programmes de très grand séquençage du génome humain permettent aujourd'hui de disposer d'un nombre significatif de nouvelles séquences rétrovirales. L'usage de logiciels de traitement de données permet d'identifier et d'analyser ces gènes. Dans ce contexte une recherche systématique portant sur l'ensemble des informations disponibles à ce jour a été engagée afin d'identifier de nouvelles séquences

rétrovirales endogènes humaines en fonction de certains critères d'analyse :

- présence de longs cadres de lecture ouverts conservés au cours de l'évolution de l'organisme hôte et pouvant laisser envisager une fonction biologique.

- analogie avec des séquences déjà caractérisées en dehors ou dans le
5 domaine des rétrovirus,

- localisation dans des régions de susceptibilité pour certaines pathologies ou à proximité de gènes essentiels, par exemple dans les domaine du cancer, des régulation du système immunitaire ou dans certaines neuropathologies.

Les recherches effectuées par les Inventeurs, dans des bases de données de séquences leur ont permis d'identifier un ensemble de séquences ou de motifs
10 rétroviraux endogènes dont l'expression normale ou pathologique peut favoriser ou perturber un effet protecteur vis-à-vis de processus pathologiques, ou intervenir dans le déclenchement ou l'aggravation de pathologies.

La présente invention a pour objet un fragment d'acide nucléique purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une
15 séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70 % sur plus de 600 nucléotides pour
20 les domaines de type *env*.

On entend par séquence homologue, aussi bien une séquence qui présente une identité complète ou partielle avec la séquence SEQ ID NO:1 précitée qu'une séquence qui présente une similarité partielle avec ladite séquence SEQ ID NO:1.

Selon un mode de réalisation avantageux dudit fragment, il présente à la fois des motifs rétroviraux correspondant à un domaine *env* et répondant à la séquence SEQ ID NO:1 et des motifs rétroviraux correspondant à un domaine *gag* et répondant à la séquence SEQ ID NO:2 ou à une séquence présentant un niveau d'homologie supérieur ou égal à 80 % sur plus de 190 nucléotides ou supérieur ou
30 égal à 70 % sur plus de 600 nucléotides pour les domaines de type *env* et un niveau d'homologie supérieur ou égal à 90 % sur plus de 700 nucléotides ou supérieur ou

égal à 70 % sur plus de 1200 nucléotides pour les domaines de type *gag*, lesquels motifs ne présentent aucune insertion ou délétion de plus de 200 nucléotides.

Lesdits fragments constituent une nouvelle famille de séquences rétrovirales endogènes humaines (famille HERV-7q) qui présente une homologie de
5 séquence avec les rétrovirus MSRV, tels que décrits dans la Demande Internationale WO 97/06260 ; lesdits fragments selon la présente invention présentent :

- deux motifs nucléotidiques répétés de 711 pb (figure 3), présentant des signaux caractéristiques identifiés dans des LTRs (*Long Terminal Repeats*) : promoteurs de transcription de type boîtes TATAA ou CCAAT. Ces domaines répétés
10 encadrent trois motifs déduits de type-*gag*, *pol* et *env* (figure 2).

- un motif de type *env* (positions 6965 nt - 9550 nt sur la séquence SEQ ID NO :3 ou sur la figure 1) qui contient un long cadre de lecture ouvert de 1620 nucléotides (positions 7874-9493 de la séquence ID NO:3 et figure 1), codant pour une protéine de séquence inédite de 540 acides aminés appelée envérine (figure 4 et
15 SEQ ID NO:26) et fragment souligné de la figure 18. On retrouve à l'intérieur du domaine trans-membranaire de ce domaine *env*, un motif peptidique de type CKS-25/CKS-17 (figure 5), reconnu pour présenter des fonctions immunosuppressives sur les cellules lymphocytaires hôtes (M. Mitani et coll., 1987, *Proc. Natl. Acad. Sci. USA*, **84**, 237-240). Un domaine de type doigt de zinc (*zinc-finger*) $\text{HX}_3\text{HX}_{22}$
20 CX_2C (Kulkolski et coll., 1992, *Mol. Cell. Biol.*, **12**, 2331-2338), que l'on retrouve dans des domaines de type intégrase est identifié dans un autre cadre de lecture. Ce domaine *env* particulier signe la caractéristique de nouveaux motifs rétroviraux endogènes.

- le motif (positions 3065 nt - 4390 nt sur la séquence SEQ ID NO:3)
25 de type-*gag* codant pour des motifs protéiques selon la figure 6 (SEQ ID NO:58) (positions 3118-4198 de la SEQ ID NO:3) a été identifié grâce à des analogies avec des domaines *gag* connus. On retrouve, par exemple, la région d'homologie majeure $\text{QX}_3\text{EX}_7\text{R}$ (Benit et coll., 1997, *J. Virol.*, **71**, 5652-5657). Le motif de fixation des acides nucléiques $\text{CX}_2\text{CX}_3\text{HX}_4\text{C}$, situé en position C-terminale, est identifié dans un
30 autre cadre de lecture (Covey et coll., 1986, *Nucleic Acids Res.*, **14**, 623-633). En amont du domaine *gag* on détecte un motif de 182 nucléotides répété deux fois (figure

1).

- le domaine *pol* présente les consensus classiques d'une région *pol* de rétrovirus au niveau des domaines protéase, transcriptase reverse et RNase H. On retrouve dans *pol* un motif proche du consensus **LLDTGA** (Weber et coll., 1988, Science, 243, 928-931). Les motifs **D** et **AF**, **LPQ** et **SP**, et **YVDD** (Xiong et Eickbush, 1990, EMBO J., 9, 3353-3362), sont respectivement retrouvés dans les 3°, 4° et 5° boîtes d'homologie. Les motifs **YTDGSS** et **TDS** sont présents dans la région de la RNase H,

- les régions *gag* et *pol* pourraient être considérées comme jointives avec un passage de la région *gag* à la région *pol* par un décalage du cadre de lecture.

La présente invention englobe les séquences appartenant à la famille HERV-7q telle que définie ci-dessus (présence de la séquence SEQ ID NO:1 ou d'une séquence homologue ou présence à la fois des séquences SEQ ID NO:1 et SEQ ID NO:2) et notamment les séquences SEQ ID NO:3-22, 28 et 61 ; elle englobe également les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes ainsi que les fragments issus des régions codantes des séquences précédentes correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires. (SEQ ID NO :37-57, 59-60 et 121-122).

Ces différents fragments peuvent avantageusement être utilisés comme amorces ou comme sondes (réactifs A) ; ils s'hybrident spécifiquement dans des conditions de forte stringence à une séquence de la famille HERV-7q.

Parmi ces fragments, on peut citer, de préférence les fragments suivants:

- un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865 de la SEQ ID NO:3 ;

Amorces et sondes spécifiques de la région *gag*

- une amorce G1F, sens, localisée dans la région amont du domaine *gag* de HERV-7q : 5' GGACCATAGAGGACACTCCAGGACTA 3' (SEQ ID NO:37);

- une amorce G1R, anti-sens, localisée dans la région 3' terminale du

domaine *gag* : 5' CCTCAGTCCTGCTGCTGGATCATCT 3' (SEQ ID NO:38)

- le fragment de 1505 nt amplifié par le couple G1F-G1R est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR ;

- 5 - une amorce G2F, sens nichée : (SEQ ID NO:39)
 5' CCTCCAAGCAGTGGGAGGAAGAGAATT 3'
 - une amorce G2R, anti-sens nichée : (SEQ ID NO:40)
 5' CCTTCCCTGTGTTATTGTGGACATCATT 3'
 - une amorce G4F, sens nichée : (SEQ ID NO:41)
 10 5' GGAAGAAGTCTATGAATTATTCAATGATGT 3'
 - une amorce G3F, sens nichée: (SEQ ID NO:42)
 5' GGGACACAGAATCAGAACATGGAGATT 3'
 - une amorce G4R, anti-sens nichée : (SEQ ID NO:43)
 5' GCCTTCAGAAGAGTCAGGTGACAGAGA 3'
 15 - une amorce G5R, anti-sens nichée : (SEQ ID NO:44)
 5' GAGCCTCCAAAGTCCACTTGCCTGA 3'

Amorces et sondes spécifiques de la région *env*

- une amorce E1F, sens : (SEQ ID NO:45)
 5' GATTTCAGTATCTACTAGTCTGGGTAGAT 3'
 20 - une amorce E1R, anti-sens : (SEQ ID NO:46)
 5' CTAGGAAATCCAGCTAGTCCTGTCTCA 3'
 - le fragment de 2529 nt amplifié par le couple d'amorces E1F-E1R,
 est utilisé afin de générer les sondes aptes à hybrider les différents produits
 d'amplification des PCR.
 25 - une amorce E2F, sens : (SEQ ID NO:47)
 5' CCAAGACAGCCAACTTAGTTGCAGACAT 3'
 - une amorce E2R, antisens : (SEQ ID NO:48)
 5' GGACGCTGCATTCTCCATAGAAACTCTT 3'
 - une amorce E3F, sens : (SEQ ID NO:49)
 30 5' GCAATACTACATACACAACCAACTCCCAA 3'

- une amorce E3R, anti-sens : (SEQ ID NO:50)
5' GGGGGAGGCATATCCAACAGTTAGTA 3'
- une amorce E4F, sens : (SEQ ID NO:51)
5' CCATCTACACTGAACAAGATTTATACACTT 3'
- 5 - une amorce E4R, anti-sens : (SEQ ID NO:52)
5' AATGCCAGTACCTAGTGCACCTAGCACT 3'
- une amorce E5F, sens : (SEQ ID NO:53)
5' CGAATACAACGTAGAGCAGAGGAGCTTCGAA 3'
- une amorce E6F, sens : (SEQ ID NO:54)
10 5' AGCCCAAGATGCAGTCCAAGACTAAGAT 3'
- une amorce E5R : (SEQ ID NO:55)
5' GCGTAGTAGAGGTTGTGCAGCTGAGAT 3'
- une amorce ExF : (SEQ ID NO:56)
CCCTTACCAAGAGTTTCTATGGAGAAT
- 15 - une amorce ExR : (SEQ ID NO:57)
ACCGCTCTAACTGCTTCCTGCTGAATT

Tous les oligonucléotides sont conçus pour pouvoir générer une amorce sens et une amorce anti-sens par un décalage de la séquence de l'amorce de référence de 1 à 7 nucléotides vers le côté 5' ou vers le côté 3': la modification de la

20 séquence peut entraîner une modification de la taille de l'amorce de 1 à 7 nucléotides selon les cas. Les amorces choisies peuvent être optimisées selon les cas par un raccourcissement ou un allongement portant sur 1 à 9 nucléotides.

De manière préférée, l'hybridation, le clonage, le sous-clonage, l'obtention, la préparation et l'analyse des acides nucléiques, des peptides et des anti-

25 corps, le séquençage des acides nucléiques et des peptides, l'hybridation *in situ* et l'immunohistochimie sont réalisés dans les conditions décrites dans les ouvrages suivants :

- Current Protocols in Molecular Biology. Eds. F.M Ausubel, R. Brent & R.E Kingston et coll. Green Publishing associates and Wiley Interscience.
- 30 - Molecular Cloning: a laboratory manual. Eds. J. Sambrook, E.F. Fritsch & T. Maniatis. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.

- The Practical Approach series. Eds. D. Rickwood & B.D. Ames. IRL Press and Oxford University Press. En particulier, antibodies I & II; DNA cloning I, II, III; Nucleic acid and protein sequence analysis; Nucleic acid hybridization; Nucleic acid sequencing ; Oligonucleotide synthesis; Protein purification applications;
- 5 Protein purification methods; Protein sequencing; Transcription and translation; Gels electrophoresis of nucleic acids; Gels electrophoresis of proteins; Genome analysis; HPLC of macromolecules; Human genetic diseases; Microcomputing in biology; Molecular neurobiology; Mutagenicity testing; Essential molecular biology I & II.

- Proteome research: New frontiers in functional genomics. Eds
- 10 M.R. Wilkins & coll.. Springer.

La séquence rétrovirale endogène humaine (SEQ ID NO:3). située sur le bras long du chromosome 7 correspond à la séquence HERV-7q ; elle présente 10,5 kb (fig. 1 et 2) et répond aux critères précédemment définis.

- La recherche de domaines présentant des similitudes, tout ou partie,
- 15 avec les régions *gag* et *env* de HERV-7q a abouti à l'identification de nouvelles séquences rétrovirales endogènes. Ces séquences peuvent présenter la structure d'un rétrovirus endogène complet comme la séquence rétrovirale endogène située à proximité du gène des sous-unités alpha et delta du récepteur des cellules-T, et dénommée en conséquence HERV-TcR ; à titre d'exemple la figure 7 montre la comparaison des
- 20 alignements nucléiques des domaines *gag* respectifs de HERV-7q et HERV-TcR (séquence HG12, SEQ ID NO:19). On trouve aussi des structures rétrovirales partielles. Ces domaines rétroviraux similaires à HERV-7q sont identifiées dans des séquences nucléiques indépendantes comme le montre leur localisation chromosomique. Des motifs nucléiques (appelés ici, HEx ou HGx et respectivement analogues à des domaines de type *env* ou *gag*) ressemblant aux domaines *env* ou *gag* de
- 25 HERV-7q ont été retrouvés, à l'aide des banques de données précitées :

- HE2 : chromosome 17 (SEQ ID NO:4),
- HE3 et HG3: chromosome 6 (SEQ ID NO:5 et 6),
- HE4 : chromosome X (SEQ ID NO:7),
- 30 - HE5 : chromosome X q22 (SEQ ID NO:8),
- HE6 et HG6 : chromosome 1 q23.3-q24.3 (SEQ ID NO:9 et 10),

- HE7 : chromosome 7 p15 (SEQ ID NO:11),
- HE8 et HG8 : chromosome 19 (SEQ ID NO:12 et 13),
- HE9 : chromosome X (SEQ ID NO:14),
- HE10 : chromosome X q13.1-21.1 (SEQ ID NO:15),
- 5 - HE11 et HG11 : chromosome 7 q21-22 (SEQ ID NO:16 et 17),
- HE12 et HG12, dans HERV-TcR : chromosome 14 q11.2 (SEQ ID NO:18 et 19),
- HE13 (SEQ ID NO:61) : chromosome 6 q24.1-24.3

La présente invention englobe également les fragments codants et
10 non codants pour tout ou partie de l'envérine comprenant au moins 14 nucléotides et notamment les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première méthionine.

Ces fragments comprennent en particulier une zone critique où deux
15 insertions de 12 nucléotides ont été caractérisées :

- une première insertion a été identifiée (séquence A), chez des individus de 2 groupes (malades et témoins). Cette insertion située entre les acides aminé 487 et 488, permet d'insérer le térapeptide VLQM. Une analyse comparative montre que cette insertion est identifiée dans une région homologue située dans la séquence
20 HE13, appartenant à la famille HERV-7q. L'amplification de la séquence de type HE13, pourrait indiquer qu'il existe une altération de la séquence de l'envérine de HERV-7q, ce qui favoriserait l'amplification de la séquence contenue dans HE13. Cette observation permet aussi d'utiliser cette insertion comme élément spécifique d'amplification de séquences de type HE13.

25 Une deuxième insertion (séquence B) a été identifiée chez un patient présentant une SEP. L'insertion de 12 nucléotides est située au niveau de l'acide aminé 495 et code pour le térapeptide MQSM. Il est remarquable de constater que cette insertion est aussi identifiée dans une région homologue située dans HE13.

Séquence A: TAACTACAAATGGTTCTTCAAATGGAGCCCA
30 (SEQ ID NO:59)

Séquence B: GATGCAGTCCAAGATTGCAGTCCATGACTAAGA
(SEQ ID NO:60).

Ces observations mettent en évidence des modifications de la séquence de l'envérine de type HERV-7q qui constituent la base d'une stratégie de
5 détection par amplification spécifique d'allèles (AS-PCR), permettant de détecter ces différences dans une population et qui pourraient correspondre, soit à une mutation/délétion associée à une certaine susceptibilité, soit à un polymorphisme, soit à une mutation/délétion associée à une pathologie comme la sclérose en plaques.

Les alignements des domaines *env* (fig. 8) et *gag* (fig. 9) explicitent
10 les niveaux d'homologie observés entre les séquences décrites ci-dessus et les séquences homologues dans HERV-7q. Les analogies peuvent s'étendre aux motifs rétroviraux flanquants.

Une analyse des séquences étiquettes disponibles dans les banques de données montre que des transcrits appartenant à certains des membres de cette
15 famille, en particulier HERV-7q, s'expriment essentiellement dans des tissus d'origine fœtale ou placentaire.

Des séquences polypeptidiques générées par ces transcrits peuvent donc être potentiellement produites et des fonctions ou activités biologiques peuvent être envisagées, par analogie avec des polypeptides biologiquement actifs d'origine
20 virale ou rétrovirale ; par exemple, les motifs peptidiques de type CKS-17 (Haraguchi et al., PNAS, 1995, 92, 5568-5571) (fig. 5) ou CKS-25 (Huang S.S et Huang J.S, J. Biol. Chem. 1998, 273, 4815-4818), qui présentent des fonctions immunomodulatrices sur les cellules lymphocytaires hôtes. Les différences de séquence observées et d'éventuelles modifications normales ou pathologiques, sont en particulier, à l'origine
25 d'une modulation de la fonction.

HERV-7q représente le paradigme de la nouvelle famille de séquences rétrovirales endogènes humaines ou de motifs rétroviraux endogènes.

HERV-7q et certaines des séquences rétrovirales endogènes appartenant à sa famille, présentent un domaine de type *pol* analogue à des séquences rétro-
30 virales de type *pol* comme par exemple la région *pol* identifiée dans le rétrovirus MSRV associé à la sclérose en plaques et décrit par H. Perron et al. (1997, *Proc. Natl.*

Acad. Sci. USA, 94, 7583-7588 ; Demande Internationale PCT WO 97/06260).

Toutefois, les séquences selon la présente invention se distinguent des séquences rétrovirales exogènes infectieuses analogues à MSR₁V antérieurement décrites en ce que les séquences *gag* et *env*, selon l'invention sont significativement
5 différentes selon les critères précédemment définis et en fonction de certaines caractéristiques spécifiques, par exemple le long cadre de lecture ouvert du domaine *env* de HERV-7q ; elles seraient à même de permettre de signer une pathologie lorsqu'elles présentent des insertions, des délétions, des décalages de cadre de lecture ou des mutations.

10 En effet, les différences observées entre les séquences humaines de type HERV-7q, qui sont isolées d'individus réputés normaux et les séquences issues de certains échantillons d'origine pathologique, ne sont pas distribuées au hasard. Des comparaisons menées entre la région *gag* provenant de particules rétrovirales infectieuses (N° d'accèsion EMBL: A60168, A60200, A60201, A60171...) et la séquence
15 *gag* correspondante de HERV-7q (fig. 9), permettent d'observer que les mutations affectent préférentiellement des codons non-sens. Par exemple, deux codons non-sens dans HERV-7q sont remplacés par un codon arginine dans A60200, ce qui permet d'obtenir une séquence déduite de 109 acides aminés pour HERV-7q et de 166 acides aminés pour A60200. Les changements de base permettent en conséquence de prolonger le cadre de lecture et de coder potentiellement pour des structures polypeptidiques
20 de plus grande taille (figure 10).

De même, une séquence de type *env* provenant de particules rétrovirales infectieuses, présente une analogie significative avec le domaine *env* de HERV-7q (figure 11). Ces analogies marquées entre séquences rétrovirales exogènes
25 et endogènes pourraient être à l'origine du déclenchement ou de l'aggravation de certains processus pathologiques, en particulier de certaines maladies auto-immunes, comme la sclérose en plaques. A cet égard, on peut remarquer que certaines des séquences rétrovirales endogènes décrites dans l'invention se situent à proximité ou dans des régions réputées présenter une susceptibilité pour la sclérose en plaques : par
30 exemple HERV-7q et la région 7q21-22 du chromosome 7, de même pour HE12 et HG12 dans HERV-TcR et la région du gène codant pour les chaînes alpha et delta du

récepteur des cellules-T, HE2 et le chromosome 17, ou HE3, HE13 et HG3 et le chromosome 6, par exemple, les séquences HE11 et HG11, autour de la région 7q 21-22 ou encore HE4, HE5, HE6, HE9, HE10 ou HG10 sur le chromosome X. Ces séquences seraient donc à même de fournir des moyens de localisation ou d'identification des gènes de prédisposition.

On n'observe aucune homologie significative avec des séquences rétrovirales endogènes déjà décrites, par contre, on peut relever une homologie limitée, permettant d'identifier une structure générale de domaine *env* : toutefois, ladite homologie est inférieure aux critères définis selon l'invention entre les domaines *env* de la séquence HERV-7q (SEQ ID NO :1) et de la séquence HERV-9 (figure 12). La figure 11 montre des homologies étendues entre la séquence HERV-7q avec une séquence rétrovirale exogène (N° d'accension EMBL : A60170).

Les séquences rétrovirales endogènes humaines appartenant à la famille de HERV-7q, peuvent protéger contre des agressions liées à l'environnement ou constituer un bénéfice pour l'individu. Cet effet bénéfique pourrait être une des raisons possibles de la pression de sélection exercée sur certaines de ces séquences et du caractère potentiellement fonctionnel des structures protéiques déduites identifiées : par exemple le long cadre de lecture ouvert apte à coder pour une nouvelle protéine et correspondant au domaine *env* de HERV-7q.

Les séquences rétrovirales endogènes humaines appartenant à la famille de HERV-7q pourraient être associées par exemple, à des pathologies en relation avec les processus liés au cancer, aux neuropathologies à composante auto-immune ou à tout autre processus pathologique en association ou non avec des virus ou rétrovirus endogènes ou exogènes. Leur action pourrait porter sur la déclaration, l'aggravation, la modification du calendrier d'apparition ou encore la protection vis à vis de la maladie.

Dans le contexte d'application à des pathologies autoimmunes (comme par exemple le lupus, le syndrome de Sjögren, la polyarthrite rhumatoïde, la sclérose en plaques...), on peut relever des analogies significatives entre les motifs rétroviraux endogènes identifiés et des motifs retrouvés dans des structures rétrovirales caractérisées chez des patients présentant des pathologies autoimmunes comme la sclérose en plaques : par exemple des fragments de domaine *gag* (récemment dispo-

nibles dans les banques de données) provenant de particules rétrovirales infectieuses ou encore la séquence complète du domaine *pol* correspondant au virus MSRV associé à la sclérose en plaques. Ces motifs rétroviraux possèdent des analogies significatives avec les séquences endogènes homologues de type HERV-7q, ce qui permet
5 d'envisager une association directe ou indirecte avec des processus pathologiques, dont la sclérose en plaques, en association ou non avec MSRV.

L'intérêt de ces séquences dépasse le cadre des maladies auto-immunes. En dehors de l'importance générale des motifs rétroviraux dans le déclenchement ou l'aggravation d'un processus tumoral, bien montré en particulier dans les
10 modèles murins (H. Fan dans *The retroviridae*, 1994, ed. J.A. Levy, Plenum. New York, p. 313-353), ces séquences pourraient se retrouver à proximité ou au sein de gènes importants et en altérer l'expression : par exemple HERV-TcR et les gènes des sous-unités alpha et delta du récepteur des cellules T impliquées dans des perturbations de la fonction immunitaire.

15 La présente invention englobe, en outre, l'utilisation de séquences associées aux séquences de la famille HERV-7q pour la détection et/ou le pronostic de différentes maladies auto-immunes (neuropathologies, en particulier) ; ces séquences codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération (polyadénylation, épissage alternatif) est associée à l'expression normale
20 ou pathologique ou à la régulation/dérégulation des motifs appartenant à la famille HERV-7q et correspondent à des transcrits ou des ADNc des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes ou encadrant des séquences rétrovirales de la famille HERV-7q.

On entend par région flanquante, toute région située à proximité
25 (incluse dans ou incluant) une séquence rétrovirale endogène appartenant à la famille HERV-7q, telle que définie ci-dessus, jusque et y compris les gènes immédiatement contigus et/ou situés à une distance ne pouvant excéder 120 kb.

Les Inventeurs ont maintenant trouvé que la présence des séquences rétrovirales telles que définies ci-dessus, perturbent l'expression ou altèrent la struc-
30 ture des séquences flanquantes définies ci-après.

Les transcrits desdites séquences flanquantes (et leurs fragments,

notamment ceux soulignés ou en italique dans les figures 14-16, 22-26, sont définies ci-après :

- à 1021 pb en amont de HERV-7q, on identifie une séquence rétrovirale endogène appelée RH7 (SEQ ID NO:62 et figure 22) ; cette séquence est située en 5' de la séquence HERV-7q ; dans la figure 22, la partie en italique correspond au début de la séquence HERV-7q ; la séquence RH7 est soulignée : deux sites de polyadénylation putatifs sont en gras. Cette séquence SEQ ID NO:62 présente une homologie significative, sur plus de 6 kb, avec des séquences rétrovirales endogènes de type RGH (figure 13). Des séquences appartenant à cette famille s'expriment en particulier chez des patients présentant une arthrose rhumatoïde (Nakagawa et coll., (1997), Arthritis. Rheum., 40, 627- 638). La présente invention inclut également des fragments de la séquence SEQ ID NO:62, comprenant entre 14 et 50 nucléotides (utilisation comme amorces), de préférence entre 14 et 25 nucléotides ou au moins 25 nucléotides (utilisation comme sonde), lesquels fragments présentent les caractéristiques suivantes : les 4 nucléotides de l'extrémité 3' sont différents des motifs correspondant de la séquence RGH2 (séquence du bas dans la figure 13, n° accession à GenBank : D110 18).

- à moins de 9 kb en amont de HERV-7q, on identifie la séquence RAM75 (SEQ ID NO:63 et figure 14) contenant les 24 exons codants (qui couvrent près de 41 kb), du gène de l'ATPase péroxysomale PEX1. PEX1, en association avec PEX6 est responsable de l'importation des protéines péroxysomales et de la stabilisation du récepteur PEX5. Une perturbation/altération affectant PEX1 est responsable de diverses neuropathologies comme le syndrome de Zellweger, l'adrénoleucodystrophie néonatale et la forme infantile de la maladie de Refsum (Reuber et coll., (1997), Nature Genet., 17, 445- 448). On peut rappeler que la fonction principale des péroxysomes est associée au métabolisme des acides gras, en particulier par des processus de β -oxydation. Une altération du gène identifié dans la séquence RAM75 ou de son expression, par modification de la fonction des régions 5' et 3' régulatrices ou encore par modification des épissages ou des processus de polyadénylation, en particulier sous l'influence de motifs rétroviraux voisins, seraient à même de perturber l'expression ou la structure de l'ATPase et par conséquent de perturber une des

fonctions péroxysoniales, en particulier le métabolisme des lipides, en particulier myéliniques, avec des conséquences pour certaines pathologies, dont des neuropathologies, comme la sclérose en plaques ; les parties soulignées (figure 14) correspondent aux 24 exons codants.

5 La présente invention inclut également les fragments de la séquence SEQ ID NO:63, inclus dans les 24 exons codants précités et comprenant au moins 14 nucléotides.

L'analyse du profil d'expression (transcrits et protéines) de la séquence RAM75 (SEQ ID NO:63) est un bon indicateur du diagnostic différentiel des
10 neuropathologies à composante auto-immune.

A la figure 14, les exons codants sont soulignés. Les codons d'initiation et non-sens ainsi que les sites putatifs de polyadénylation sont en gras et soulignés.

- à 0.7 kb en aval de la séquence HERV-7q et sur près de 17 kb
15 (SEQ ID NO:64 et figure 15), on identifie la séquence nucléotidique RAV73, où l'on détecte des séquences étiquettes et des exons potentiels aptes à produire une ou plusieurs séquences polypeptidiques ; l'invention inclut également des fragments de cette séquence SEQ ID NO:64 compris dans les séquences étiquettes et les exons potentiels tels qu'ils apparaissent (parties soulignées) à la figure 15, lesquels
20 fragments comportent au moins 14 nucléotides.

- à 120 kb en amont de la séquence HG3, et sur 15 kb, se trouve la séquence nucléotidique RBP3 (SEQ ID NO:65 et figure 23), qui couvre l'extrémité 3'du gène codant pour un facteur de transcription de la famille Blimp-1 (SEQ ID NO:119 et figure 25), une protéine de 789 acides aminés qui est un répresseur de
25 l'expression du gène de l'interféron-béta (Keller et Maniatis, Genes Dev., (1991), 5, 868-879), qui est déjà associé à certaines pathologies malignes (Mock et coll., Genomics, (1996), 37, 24-28), et qui pourrait jouer un rôle dans la différenciation et la pathogenèse des cellules B. L'intérêt de l'association possible de la séquence rétrovirale endogène contenant les motifs HG3 et HE3 et de Blimp-1 est multiple, dans le
30 cas de pathologies, et en particulier la sclérose en plaques. Blimp-1 agit en particulier sur les cellules B dont on connaît la contribution dans les processus inflammatoires

associés à la sclérose en plaques. Blimp-1 est capable de bloquer l'induction virale du promoteur $INF\beta$ dont on connaît l'aptitude à réduire la fréquence des poussées et la progression lésionnelle chez des patients atteints de SEP. Une perturbation de l'expression ou de la structure de Blimp-1, en relation avec un élément rétroviral de type HERV-7q, est associée en conséquence à des neuropathologies ou à des maladies à caractère auto-immun, comme la sclérose en plaques ; cette séquence nucléotidique RBP3 (SEQ ID NO:65) contient des motifs nucléotidiques identifiés dans la séquence nucléique codant pour le gène Blimp-1 ; l'invention inclut aussi la détection des séquences ARNm de la protéine Blimp-1 (SEQ ID NO:119).

10 - la séquence rétrovirale endogène de type HERV-7q, contenant HE3 et HG3, se trouve située dans la région HI3 correspondant à un intron s'étendant sur plus de 46 kb (SEQ ID NO:66), d'un gène qui pourrait coder pour l'analogue d'APS (figure 24), une protéine de 275 acides aminés spécifique d'apoptose, surexprimée dans différents cellules en culture après déclenchement d'un processus apoptotique
15 (Hammond et coll., FEBS Lett., (1998), 425, 391- 395). L'intron se situe au niveau de l'acide aminé 231 d'APS. L'extrémité de HE3 est à plus de 12 kb de l'extrémité 5' de l'intron, alors que HG3 est situé à plus de 28 kb de l'extrémité 3' de l'intron. Des processus apoptotiques sont associés à la sclérose en plaques. En particulier, il a été décrit un processus apoptotique affectant des astrocytes et des oligodendrocytes en
20 présence d'une fraction purifiée de liquide céphalo-rachidien de patients atteints de sclérose en plaques (Ménard et coll., J. Neurol. Sci., (1998) , 154, 209- 221).

Enfin, il faut souligner que la région nucléique contenant HE3, HG3, HI3 et RBP3 est localisée au niveau du bras court du chromosome 6, en 6p21, qui est une région proposée de susceptibilité à la sclérose en plaques (The Multiple Sclerosis
25 Genetic Group, Nature Genet., (1996), 13, 469- 472).

L'interaction entre les séquences de type HERV-7q et les séquences flanquantes et l'importance de l'établissement d'un profil d'expression incluant une ou plusieurs des séquences précitées pour établir un diagnostic différentiel d'une neuropathologie apparaît encore plus du fait que l'on observe que les séquences HG12 et
30 HE12 sont situées dans une région intronique du gène codant pour les sous-unités alpha et delta des récepteurs des cellules T. Les récepteurs des cellules T sont impli-

qués dans les processus de régulation immunitaire et leur influence a été proposée dans le cas de maladies auto-immunes, dont la sclérose en plaques.

L'invention a également pour objet les transcrits générés à partir des séquences précitées ainsi que celles présentant éventuellement des modifications avec
5 les séquences de référence décrites dans l'invention lorsqu'ils sont exprimés chez certains patients.

En effet, les systèmes de régulation de l'expression des protéines rétrovirales de HERV-7q, qui sont présents dans les motifs de type LTR, pourraient influencer l'expression de gènes situés dans le voisinage chromosomique proche ou
10 éloigné et induire des perturbations à caractère immunologique et/ou neurologique. Par exemple la séquence rétrovirale endogène HERV-TcR, se trouve à proximité immédiate des gènes des sous-unités alpha et delta du récepteur des cellules-T précédemment décrit. Les motifs de type LTR pourraient aussi coder pour des superanti-
gènes (Acha-Orbea et Palmer, 1991, *Immunol. Today*, 12, 356-361). D'une manière
15 générale des protéines rétrovirales de type HERV-7q ou apparenté, ou leurs formes tronquées ou partielles pourraient être impliquées dans des phénomènes de cytotoxicité ou de superantégenité, comme par exemple celles issues du long cadre de lecture ouvert identifié dans le domaine *env* (figure 4).

Des séquences du type des LTR 5' et 3' de HERV-7q, fortement
20 conservées sont concernées par de tels effets régulateurs. A titre d'exemple on décrit LTX, une séquence comparable à celle d'une LTR de HERV-7q (SEQ ID NO:67 et figure 16), et qui se trouve au cœur d'un intron de plus de 49 kb, mais à 2 kb du site 5' donneur, du gène de FMR2 associé au X-fragile et codant pour une protéine de 1311 acides aminés (figure 26). Les LTR modulent l'épissage alternatif (Kapitonov et
25 Jurka, (1999), *J. Mol. Evol.*, 48, 248- 251), l'expression du gène, la fixation sur des protéines nucléaires (Akopov et coll., (1998), *FEBS Lett.*, 421, 229- 233), ou permettent l'obtention d'un signal de polyadénylation alternatif (Goodchild et coll., (1992), *Gene*, 121, 287- 294).

D'une manière générale, on peut remarquer l'existence de plusieurs
30 séquences rétrovirales endogènes de type HERV-7q (HE4, HE5, HE9, HE10), situées au niveau du chromosome X qui représente le chromosome associé au plus grand

nombre de pathologies.

A cet égard, on peut relever que des motifs rétroviraux issus de régions-défectives sont aptes à présenter des fonctions biologiques: par exemple, la protéine d'enveloppe p15E issue de motifs rétroviraux défectifs, possède une activité
5 anti-inflammatoire et immunosuppressive (Snyderman et Ciancolo, 1984, *Immunol. Today*, 5, 240-244).

Ces structures sont vraisemblablement à même de provoquer des brèches ou d'amplifier des dérégulations dans les processus de défense immunitaire. Certains des motifs des domaines *gag*, *env* et de type LTR peuvent être associés à une
10 fonction particulière ou peuvent contribuer à la fonction normale ou pathologique des domaines flanquants tels que définis ci-dessus (SEQ ID NO:62-67). Des recombinaisons avec un élément d'origine exogène, rétroviral ou non, peut donner lieu à la production de motifs nucléiques ou protéiques qui pourraient soit protéger, soit déclencher, ou favoriser ou aggraver une pathologie. De même, une structure rétro-
15 virale contenant des éléments rétroviraux endogènes selon l'invention seraient à même de provoquer un processus pathologique après passage par un cycle transitoire exogène puis réintégration dans une région sensible ou critique du génome humain.

On peut ainsi obtenir des profils d'expression (transcrits et éventuellement protéines) qui correspondent aux neuropathologies précitées.

20 De même, la combinaison de motifs appartenant à la famille de HERV-7q, ou d'éléments induits par des motifs appartenant à la famille de HERV-7q, avec des motifs d'origine ou induits de manière exogène seraient à même de pouvoir déclencher, ou aggraver un processus pathologique ou au contraire de favoriser une protection ou une rémission partielle ou une guérison totale et définitive.

25 La détection rendue possible des domaines de type HERV-7q, suggère des applications possibles à la fois au niveau prophylactique, du pronostic et du diagnostic: par exemple des approches immunologiques ou d'amplification génique permettant de comparer des individus normaux servant de référence avec des patients, seraient à même de favoriser le dépistage, d'améliorer la détection précoce de
30 la déclaration de la maladie et/ou de suivre l'évolution d'une pathologie chez des patients pouvant présenter une susceptibilité ou ayant déclaré la maladie ou encore

chez des individus considérés comme normaux, selon les critères cliniques actuels.

Les sondes nucléiques et immunologiques spécifiques, telles que définies, dans la présente invention sont à même de favoriser l'identification et la détection de motifs anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier auto-immunes, au premier rang
5 desquelles la sclérose en plaques.

La présente invention a également pour objet des séquences nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs appartenant à la famille de HERV-7q, ou d'éléments induits par des motifs appartenant à la famille de HERV-7q, avec des motifs d'origine ou induits de manière
10 exogène (séquences rétrovirales exogènes); de telles séquences hybrides sont vraisemblablement à même de pouvoir déclencher, ou aggraver un processus pathologique ou au contraire de favoriser une protection ou une rémission partielle ou une guérison totale et définitive.

La présente invention a également pour objet un réactif de diagnostic pour la détection différentielle de séquences nucléiques endogènes humaines complètes ou partielles, présentant des motifs rétroviraux, sélectionnés parmi les séquences SEQ ID NO :1 et/ou SEQ ID NO :2, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1-22, 28, 37-57, 59-61 et 121-
20 122, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes, par les fragments nucléotidiques capables de définir ou d'identifier les séquences SEQ ID NO:1 et/ou SEQ ID NO:2 et toute séquence flanquante ou les chevauchant ainsi que par les fragments issus des régions codantes des séquences SEQ ID NO:1-22 et 61, correspondant à un cadre glissant
25 supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, éventuellement marquées avec un marqueur approprié ainsi que par les séquences telles que définies aux figures 18-21.

Les séquences des sondes nucléiques, ribonucléiques et oligonucléotidiques utilisées seront choisies dans les régions *env* et *gag* ou leur régions
30 flanquantes : par exemple les oligonucléotides amorces pour HERV-7q, seront choisis dans les régions situées entre les nucléotides 3065 et 4390, les nucléotides 6965 et

9550 ou les nucléotides 2502-2865 de la SEQ ID NO:3, ainsi que dans une séquence adjacente (amont ou aval) capable de permettre une amplification spécifique (figure 1).

Parmi les marqueurs appropriés, on peut citer, les isotopes radio-actifs, les enzymes, les fluorochromes, des marqueurs chimiques (biotine), les haptènes (digoxygénine) et les anticorps ou analogues de bases appropriées.

De manière préférée :

- ledit réactif est sélectionné parmi les séquences SEQ ID NO:37-57 et est apte à être utilisé comme amorce.
- 10 - ledit réactif est sélectionné parmi les séquences suivantes :
 - un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),
 - un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R),
 - 15 un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865,
 - des fragments codants ou non-codants pour tout ou partie de l'envérine, comprenant au moins 14 nucléotides et notamment les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir
 - 20 de l'acide aminé 321, à compter de la première méthionine,
- et est apte à être utilisé comme sonde.

La présente invention a également pour objet un procédé de détection rapide et différentiel des séquences nucléiques rétrovirales endogènes de type *env* ou *env* et *gag*, de leurs variants normaux ou pathologiques, par hybridation et/ou

25 amplification génique, réalisé à partir d'un échantillon biologique, lequel procédé est caractérisé en ce qu'il comprend :

- (a) une étape dans laquelle l'on met en contact un échantillon biologique à analyser avec au moins une sonde telle que définie ci-dessus et
- (b) une étape dans laquelle on détecte par tout moyen approprié, le
- 30 ou les produits résultant de l'interaction séquence nucléotidique-sonde.

Conformément audit procédé, il peut comprendre :

* préalablement à l'étape (a) :

une étape de préparation du tissu ou du liquide biologique concerné,

une étape d'extraction de l'acide nucléique à détecter, et

5 au moins un cycle d'amplification génique et

* postérieurement à l'étape (b) :

une étape de comparaison des séquences nucléiques obtenues dans ledit échantillon biologique avec les séquences rétrovirales endogènes humaines selon l'invention par tout moyen approprié et notamment par séquençage, Southern-blot, 10 coupure de restriction, SSCP ou toute autre méthode permettant d'identifier une insertion ou une délétion ou encore une simple mutation entre les différentes séquences comparées.

Conformément à l'invention, les séquences rétrovirales endogènes humaines selon l'invention sont ainsi comparées aux séquences nucléiques présentes 15 dans l'échantillon biologique à analyser et permettent la détection de séquences homologues de patients atteints de pathologies, susceptibles de mettre en jeu une modification de leur génome.

De manière avantageuse, lesdites comparaisons géniques sont menées à partir d'ADN génomique provenant d'individus témoins et de patients.

20 Une amplification génique classique par PCR sera menée à l'aide d'amorces 5' -sens et 3' -antisens encadrant ou comprenant la zone à étudier (zone *env* ou zone *gag*).

Également de manière avantageuse, les séquences des sondes nucléiques, ribonucléiques et oligonucléotidiques utilisées sont choisies dans les 25 régions *env* et *gag* ou leurs régions flanquantes : par exemple les oligonucléotides amorces pour HERV-7q, seront choisis dans les régions situées entre les nucléotides 3065 et 4390 et les nucléotides 6965 et 9550, ainsi que dans toute séquence adjacente (amont ou aval) capable de permettre une amplification spécifique (figure 1), comme précisé ci-dessus. Elles sont de préférence sélectionnées dans le groupe constitué par 30 un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces EIF et E1R).

L'étape d'amplification génique est notamment réalisée à l'aide d'une des techniques d'amplification génique suivante : amplification par la Q β -
5 réplicase, PCR, LCR, ERA, CPR ou SDA.

La présente invention a également pour objet des séquences chimères, caractérisées en ce qu'elles sont constituées par un fragment de 17 à 40 nucléotides d'une séquence flanquante telle que définie ci-dessus associée à un motif rétroviral endogène de type HERV-7q comprenant entre 17 et 40 nucléotides, tel que
10 défini ci-dessus.

La présente invention a également pour objet un procédé de détection des transcrits, tels que définis ci-dessus, caractérisé en ce qu'il comprend :

- le prélèvement des ARN messagers provenant d'échantillons biologiques (tissus, cellules, fluides biologiques) témoins et d'un échantillon analogue
15 prélevé chez des patients et

- l'analyse qualitative et/ou quantitative desdits ARNm, par hybridation *in situ*, par dot-blot, Northern-blot, RNase mapping ou RT-PCR, à l'aide d'un réactif de diagnostic tel que défini ci-dessus.

La présente invention a également pour objet une méthode de détection et/ou d'évaluation d'une sur-expression/sous-expression ou d'une modification
20 d'au moins l'une des séquences ou fragments de séquences rétrovirales endogènes de type HERV-7q et/ou de leurs séquences flanquantes associées, caractérisée en ce qu'elle comprend :

- le dépôt sur un support approprié comme par exemple un filtre de
25 nylon, une lame de verre ou leur équivalent, de l'ADNc ou son équivalent provenant de clones, de produits de PCR obtenus à partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, lesdites séquences d'ADN étant des séquences ou des fragments de séquences rétrovirales endogènes de type HERV-7q et/ou leurs séquences flanquantes, telles que
30 définies ci-dessus, constituées par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation

ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite
5 famille HERV-7q et dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb, et/ou une séquence chimère telle que définie ci-dessus.

- l'hybridation dudit support avec au moins une sonde marquée de manière adéquate obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins
10 réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, et

- la détection des hybrides formés.

Selon un mode de mise en œuvre avantageux de ladite méthode,
15 ledit transcrit ou ADNc est sélectionné dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

Selon un autre mode de mise en œuvre avantageux de ladite méthode, ledit support comprend en outre toute séquence rétrovirale endogène ou
20 exogène.

La méthode des puces à ADN (Bowtell, (1999), Nature Genet., 21, 25- 32), est utilisée pour évaluer la modification de l'expression de tout ou partie de certaines des séquences d'origine rétrovirale de type HERV-7q et des séquences flanquantes. Brièvement de l'ADN provenant de clones, de produits de PCR obtenus à
25 partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, sont déposées sur un support, comme par exemple un filtre de nylon, une lame de verre ou leur équivalent. Les séquences nucléiques déposées couvrent les différents domaines rétroviraux décrits ci-dessus, ainsi que les séquences contiguës et les gènes flanquants. Afin de détecter d'éventuels
30 processus d'épissage alternatifs, des ADN spécifiques sont synthétisés par pas de 500- 600 nucléotides avec un chevauchement de 250- 300 nucléotides de part et d'autre.

- Les épissages alternatifs déjà identifiés feront l'objet d'une synthèse spécifique. L'hybridation s'effectue à l'aide d'une sonde obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, dont la sclérose en plaques. Dans ce cas une fraction de μg et jusqu'à quelques μg d'ARNm ou jusqu'à quelques μg ou dizaines de μg d'ARN, selon la méthode utilisée et la taille de la puce d'ADN concernée, sont suffisants pour la synthèse de la sonde nucléique.
- 10 La sonde nucléique est marquée de manière adéquate, afin d'autoriser une détection ultérieure, comme par exemple par fluorescence ou par une méthode équivalente.

L'usage de sondes bi-, voire multicolores permet de préciser l'expression concertée de plusieurs gènes en parallèle, en bénéficiant de plus d'une normalisation précise. L'acquisition des résultats est effectuée automatiquement,

15 comme par exemple par un système de balayage laser ou son équivalent.

Deux types de puces à ADN sont conçues, d'une part des puces présentant un ensemble exhaustif de séquences, et d'autre part des puces à ADN spécifiques permettant un ciblage sur une application plus spécifique.

Par exemple, une séquence critique en ce qu'elle contiendrait une

20 différence portant sur une délétion, voire une mutation, est détectée à l'aide d'oligonucléotides spécifiques (Wang et coll., (1998), Science, 280, 1077- 1082). Le polymorphisme associé à une base ou à une mutation est détecté à l'aide de quatre oligonucléotides possédant une des quatre possibilités de séquence au niveau d'une base (A, C G ou T): pour chaque différence ponctuelle les 4 oligonucléotides sont

25 déposés et les intensités d'hybridation sont comparées. De plus, un épissage alternatif est détecté en utilisant des ADN correspondant à un seul exon effectif ou putatif : le gène est donc analysé exon par exon. Les puces à ADN concernent aussi, par extension, toute séquence rétrovirale endogène ou exogène, comme par exemple ERV-9, ERV-K, ERV-L, ERV-H, ERV-4, ERV-6, ERV-8, ERV-10, ERV-15, ERV-16, ERV-

30 17, ERV-18, ERV-21, ERV-24, ERV-33, ERV-34, ERV-36, ERV-40, ERV-42, ERV-MLN, ERV-FRD, ERV-FTD...), ainsi que toutes les séquences exoniques putatives

(identifiées par l'existence de séquences étiquettes et des transcrits correspondants) ou effectives, et qui sont situées de part et d'autre jusqu'à une distance de 120 kb des séquences rétrovirales endogènes de type HERV-7q.

L'étude comparative est menée entre un échantillon témoin et
5 l'échantillon à tester, dans une perspective prophylactique, diagnostique ou thérapeutique, comme par exemple: la détection précoce d'une modification de l'expression d'une des séquences, dans une cellule, un tissu ou un organisme, l'identification d'une séquence associée à une susceptibilité ou à une pathologie quelconque, le suivi de l'évolution de la pathologie, ou encore le suivi d'un traitement et l'évaluation de son
10 efficacité.

En dehors des applications déjà énoncées, l'intérêt de la méthode permet, d'une manière plus générale, de faire un bilan des variations constatées chez un individu, ce qui constitue en quelque sorte une carte d'identité, ce qui facilite une approche épidémiologique permettant d'établir de nouvelles corrélations entre un
15 profil particulier observé et une pathologie, en dehors de tout *a priori* concernant cette pathologie.

La présente invention a également pour objet un kit de détection et/ou d'évaluation d'une maladie auto-immune et notamment des neuropathologies à étiologie auto-immune, caractérisé en ce qu'il comprend outre les tampons nécessaires
20 à la mise en œuvre des procédés tels que définis ci-dessus :

- des réactifs A de diagnostic tels que définis ci-dessus, et
- des réactifs B constitués par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique
25 ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb,

30 lesquels réactifs sont de préférence fixés sur un support approprié.

Selon un mode de réalisation avantageux dudit kit, lesdits réactifs B

sont sélectionnés dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, ainsi que les séquences représentées aux figures 13-17, 22-26.

5 La présente invention a également pour objet des produits de traduction, caractérisés en ce qu'ils sont codés par une séquence nucléotidique telle que définie ci-dessus.

La présente invention a également pour objet un peptide, caractérisé en ce qu'il est susceptible d'être exprimé à l'aide d'une séquence nucléotidique sélectionnée dans le groupe constitué par les séquences SEQ ID NO:1-22, 28 et 61, telles que définies ci-dessus, selon les combinaisons offertes par l'usage des différents cadres de lecture possibles (voir également figures 18-21).

Ledit peptide englobe également les peptides ou polypeptides dérivés comprenant entre 5 et 540 aminoacides (SEQ ID NO:23-36 et SEQ ID NO:58 et leurs fragments d'au moins 5 aminoacides) et notamment un fragment de 538 aminoacides, commençant à la première méthionine de la séquence SEQ ID NO:26 (envé-
rine).

Selon un mode de réalisation avantageux desdits peptides, ils sont notamment sélectionnés parmi les séquences SEQ ID NO:23-36, 58, notamment la
20 séquence SEQ ID NO:26 et ses fragments C-terminaux, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première méthionine.

Selon un autre mode de réalisation avantageux desdits peptides, ils sont obtenus à partir des séquences nucléiques telles que définies ci-dessus, dans lesquelles au moins un codon non-sens peut être remplacé par un codon codant pour
25 l'un des aminoacides suivants : Phe (F), Leu (L), Ser (S), Tyr (Y), Cys (C), Trp (W), Gln (Q), Arg (R), Lys (K), Glu (E) ou Gly (G).

L'invention englobe ainsi les peptides déduits ou les protéines déduites correspondant à tout ou partie des séquences nucléiques décrites dans l'invention, et présentant éventuellement des modifications avec les séquences de références décrites dans l'invention, lorsqu'ils sont exprimés chez certains patients. En
30 particulier, l'invention englobe les séquences complètes ou partielles obtenues selon

les 3 cadres de lecture sens et les 3 cadres de lecture inverses et complémentaires.
(voir figures 18-21)

De manière avantageuse, l'analyse de la structure du domaine env de
HERV-7q, appelé envérine, a permis de mettre successivement en évidence:

- 5 - un peptide signal N-terminal (région 1- 21) et deux domaines
transmembranaires (région 320-340; 455-477), responsables d'interactions avec des
motifs protéiques ou lipidiques membranaires,
- un motif immunomodulateur de type CKS-17(Haraguchi et coll.,
(1995), 92, 5568- 5571)/ CKS-25. On peut noter à cet égard, la présence d'un motif
10 **RalD** à l'intérieur du peptide de type CKS-17/CKS-25 de HERV-7q et un motif **RvaD**
en position 363 qui correspondent au consensus W/RxxD, proposé pour le site actif
des TGF- β (Huang et al., J. Biol. Chem., 1997, 272, 27155- 27159). de puissants
facteurs associés à la croissance, à la différenciation et à la morphogenèse et qui sont
associés à de nombreuses pathologies humaines, comme les processus tumoraux
15 (Tang et coll., (1998), Nat. Med., 4, 802- 807) ou les maladies neurodégénératives
(Flanders et coll., (1998), Prog. Neurobiol., 54, 71- 85). Les peptides selon l'invention
contenant ces motifs peuvent avantageusement servir d'antagonistes en inhibant la
fixation des TGF- β sur leurs récepteurs naturels.

- des motifs de N-glycosylation. La glycosylation des protéines
20 d'enveloppe des rétrovirus semble être directement associée à leurs propriétés fonc-
tionnelles, par exemple en influençant le nombre des déterminants disponibles dans les
cellules-T ou en favorisant la reconnaissance des antigènes par les cellules-T. La
glycosylation pourrait jouer un rôle dans la déclaration ou l'extension d'une patholo-
gie à incidence autoimmune. Les glycosylations sont nécessaires au maintien de la
25 conformation de certains épitopes, en particulier lors de la réalisation d'une protéine
d'enveloppe recombinante à fin de mise au point d'un réactif de diagnostic et pour
favoriser l'efficacité d'un éventuel vaccin. Positions 171, 210, 216, 236, 244, 283 et
411. Nombre prévu au hasard : 3.2

- des sites de prénylation. La prénylation est un mécanisme essentiel
30 de la fixation à la membrane cellulaire et pour le ciblage de certaines protéines. Ce
processus de ciblage pourrait être essentiel pour l'élaboration d'agents thérapeutiques

spécifiques aptes à interférer dans la réalisation et la régulation du trafic de complexes cellulaires mettant en jeu des protéines impliquées dans les interactions, la croissance et les mouvements cellulaires. Positions 188 et 290. Nombre prévu au hasard : 1.8

- des sites de ciblage dans le réticulum endoplasmique. Ces sites

- 5 permettraient d'assurer le ciblage vers le réticulum endoplasmique afin d'effectuer les modifications nécessaires pour favoriser le franchissement membranaire. Positions 353 et 431. Nombre prévu au hasard : 0.2.

Par ailleurs, les Inventeurs ont montré qu'un certain nombre de peptides issus de la protéine env de HERV-7q (envérine) présentent une affinité/demi-
 10 vie élevées pour des allèles HLA de classe I. Une analyse par CADD a permis de sélectionner des peptides candidats, dont les meilleurs scores sont indiqués dans le Tableau I:

TABLEAU I

15	locali- sation	séquence	molécule HLA	score	Séquence n°
20	399	FLGEECCYYV	A-0201	7214	SEQ ID NO:68
	462	LLFGPCIFNL	A-0201	1792	SEQ ID NO:69
	189	CLPLNFRPYV	A-0201	1453	SEQ ID NO:70
	439	GLLSQWMPWI	A-0201	488	SEQ ID NO:71
	263	CLPSGIFFV	A-0201	5103	SEQ ID NO:72
25	444	WMPWILPFL	A-0201	897	SEQ ID NO:73
	252	IRWVTPPTQI	B-2705	3000	SEQ ID NO:74
	432	LRNTGPWGLL	B-2705	2000	SEQ ID NO:75
	158	LRTHTRLVSL	B-2705	2000	SEQ ID NO:76
	316	KRVPILPFVI	B-2705	1800	SEQ ID NO:77
30	25	CRCMTSSSPY	B-2705	1000	SEQ ID NO:78
	137	TRVHGTSSPY	B-2705	1000	SEQ ID NO:79
	124	AREKHVKEVI	B-2705	600	SEQ ID NO:80
	478	SRIEAVKLQM	B-2705	600	SEQ ID NO:81
	442	SQWMPWILPF	B-2705	500	SEQ ID NO:82
35	405	CYYVNQSGI	Kd	2400	SEQ ID NO:83
	346	FYYKLSQEL	Kd	2400	SEQ ID NO:84
	244	TYTTNSQCI	Kd	2400	SEQ ID NO:85
	291	SFLVPPMTI	Kd	1600	SEQ ID NO:86
	406	YYVNQSGIV	Kd	1200	SEQ ID NO:87
40	167	LFNTTLTGL	Kd	1152	SEQ ID NO:88
	463	LFGPCIFNL	Kd	960	SEQ ID NO:89
	253	RWVTPPTQI	Kd	480	SEQ ID NO:90
	449	LPFLGPLAAI	B-5102	2200	SEQ ID NO:91
	3	LPYHIFLFTV	B-5102	1210	SEQ ID NO:92

TABLEAU I (suite)

	locali- sation	séquence	molécule HLA	score	Séquence n°
5					
	331	GALGTGIGGI	B-5102	798	SEQ ID NO:93
	321	LPFVIGAGVL	B-5102	550	SEQ ID NO:94
	499	RRPLDRPAS	B-2705	600	SEQ ID NO:95
10	194	FRPYVSIPV	B-2705	600	SEQ ID NO:96
	383	RRALDLLTA	B-2705	600	SEQ ID NO:97
	39	WRMQRPGNI	B-2705	600	SEQ ID NO:98
	423	DRIQRRAEEL	B14	1800	SEQ ID NO:99
	158	LRTHTRLVSL	B14	600	SEQ ID NO:100
15	359	ERVADSLVTL	B14	540	SEQ ID NO:101
	463	LFGPCIFNLL	Kd	1658	SEQ ID NO:102
	345	QFYKLSQEL	Kd	1152	SEQ ID NO:103
	443	QWMPWILPFL	Kd	691	SEQ ID NO:104
	405	CYYVNQSGIV	Kd	500	SEQ ID NO:105
20	474	NEVSSRIEAV	Kd	480	SEQ ID NO:106
	221	GPLVSNLEI	B-5102	1320	SEQ ID NO:107
	190	LPLNFRPYV	B-5102	726	SEQ ID NO:108
	449	LPFLGPLAAI	B-5101	1144	SEQ ID NO:109
	488	EPKMQSKTKI	B-5101	968	SEQ ID NO:110
25	3	LPYHIFLFTV	B-5101	629	SEQ ID NO:111
	125	REKHVKEVI	Kk	1000	SEQ ID NO:112
	312	KPRNKRVPIL	B7	800	SEQ ID NO:113
	378	VVLQNRRAI	Db	792	SEQ ID NO:114
	377	AVVLQNRRAI	Db	660	SEQ ID NO:115
30	321	LPFVIGAGV	B-5101	629	SEQ ID NO:116
	304	DLYSYVISK	A3	540	SEQ ID NO:117
	301	TEQDLYSYVI	Kk	500	SEQ ID NO:118

Ce Tableau I indique une estimation de la demi-vie de dissociation d'un peptide de l'envérine avec un allèle du système HLA de classe I (les tables de coefficients de Parker: J. Immunol, (1994), 152, 163- 175). La localisation indique la position du premier acide aminé des peptides testés dans la séquence de l'envérine. Le code à une lettre est utilisé pour la séquence des acides aminés. Les scores autour de 500 ou supérieurs à 500 ont été retenus. A titre de comparaison, une analyse a été effectuée sur une concaténation de peptides (polypeptide de 4968 acides aminés) réputés pour fixer les molécules du complexe majeur d'histocompatibilité de classe I (Rammensee, Immunogenetics, (1995), 41, 178- 228): les dix meilleurs scores enregistrés pour des nonapeptides et le type HLA, A_0201 sont respectivement de 4984,

4047, 2406, 1267, 800, 705, 607, 591, 591 et 577.

Il ressort de ce Tableau I que certaines molécules du complexe majeur d'histocompatibilité de type I sont aptes à fixer des peptides issus de l'envérine, ainsi assimilés à des peptides d'origine virale ou tumorale, au niveau du réticulum endoplasmique. Les complexes formés au niveau du réticulum endoplasmique sont alors transportés à la surface cellulaire, ce qui entraîne la destruction de la cellule cible par les lymphocytes-T cytotoxiques. Les peptides identifiés comportent généralement 8 à 10 acides aminés. Des études ont montré que certains allèles du système HLA de classe I sont ainsi associées à certaines pathologies, en particulier à caractère auto-immun, comme HLA-B27 avec la spondylarthrite ankylosante ou HLA-B51 avec la maladie de Behçet.

Un peptide apte à fixer une molécule particulière de classe I est par conséquent apte à fonctionner comme un épitope de cellule-T.

En conséquence, la présente invention inclut également les fragments 399-471 et 244-271 de l'envérine qui regroupent avantageusement plusieurs épitopes de forte affinité pour différents haplotypes du système HLA de classe I. L'usage de tout ou partie de ces polypeptides est en conséquence apte à favoriser une augmentation du répertoire des cellules-T, en permettant une meilleure efficacité de la réponse immunitaire dans le cadre des différentes stratégies immunothérapeutiques, prophylactique ou vaccinales). Ces peptides pourront être avantageusement délivrés par exemple par l'usage, de vecteurs viraux, de particules virales ou synthétiques, de lipopeptides, d'adjuvants classiques, d'acides nucléiques nus ou adsorbés sur des particules, ou de liposomes.

Au sens de la présente invention, les peptides peuvent être chimiquement ou biochimiquement modifiés ; certaines des acides aminés peuvent être remplacés par un acide aminé analogue, selon les critères classiques d'homologies (A ou G ; S ou T ; I, L ou V ; F, Y ou W ; N ou Q ; D ou E).

La présente invention a également pour objet des compositions immunogènes ou vaccinales, pour la protection contre les maladies auto-immunes, notamment chez les sujets à risque, caractérisée en ce qu'elle comprend au moins un peptide comprenant au moins un motif de type CKS et/ou au moins un peptide

constitué par un motif présentant une affinité avec l'un des haplotypes du système HLA de classe I ou de classe II et un véhicule pharmaceutiquement acceptable.

Selon un mode de réalisation avantageux de ladite composition, ledit motif est sélectionné dans le groupe constitué par les peptides tels que définis dans le

5 Tableau I ci-dessus.

Selon un autre mode de réalisation avantageux de ladite composition, ledit peptide présente la séquence suivante :

séquence CKH: LQNRALDLLTAERGGTcIFLGEECCYYV

(SEQ ID NO:120).

10 Il est remarquable de constater au niveau de la position 380 de la protéine envérine, la contiguïté des motifs de type CKS-17 (souligné) et du peptide présentant le score le plus élevé (en gras ; voir peptide en position 399 dans le Tableau I, SEQ ID NO:68) dans la séquence CKH.

L'activation clonale des sous-groupes de lymphocytes, par exemple
15 de lymphocytes cytotoxiques, par les peptides du Tableau I et par extension leurs homologues, est bloquée par des manœuvres usuelles d'immunothérapie comme par exemple la sérothérapie et la vaccination.

L'association de deux séquences ou des séquences analogues au peptide CKH (SEQ ID NO:120), est à même d'entraîner un processus synergique dans
20 la réponse immunitaire, qui pourrait mettre en jeu des voies de signalisation et d'activation complémentaires, aptes à moduler l'activation lymphocytaire.

La vaccination concerne la production d'anticorps dirigés contre les peptides du tableau I, selon les règles de l'art et selon les méthodes de libération contrôlées par implants artificiels ou cellulaires mettant en œuvre une composition
25 telle que définie ci-dessus et par usage des manœuvres de thérapie génique, comme par exemple par expression des séquences nucléiques codant pour les peptides du Tableau I. En conséquence l'invention a également pour objet des compositions immunogènes ou vaccinale caractérisée en ce qu'elles comprennent un vecteur incluant au moins une séquence nucléique codant un peptide tel que défini dans le
30 Tableau I, éventuellement associée à une séquence codant un motif de type CKS-17.

La sérothérapie concerne l'utilisation d'anticorps neutralisants

produits à partir des peptides du Tableau I et leurs homologues.

Les produits protéiques générés par les séquences rétrovirales endogènes ou produits parallèlement peuvent avantageusement être caractérisés par des micro-méthodes d'analyse et de quantification des peptides et des protéines:

- 5 HPLC/FPLC ou équivalent, électrophorèse capillaire ou équivalent, techniques de microséquençages (méthode d'Edman ou équivalent, spectrométrie de masse...).

- L'invention a également pour objet des anticorps dirigés contre l'un ou plusieurs des peptides décrits ci-dessus et leur utilisation soit pour la mise en œuvre d'une méthode de détection *in vitro*, notamment différentielle de la présence d'une
10 telle séquence chez un individu, soit pour la préparation d'une composition apte à être utilisée en sérothérapie dans les neuropathologies à composante auto-immune.

- Lesdits anticorps sont avantageusement des anticorps polyclonaux ou monoclonaux obtenus par une réaction immunologique d'un organisme humain, mammifères, oiseaux ou autres espèces vis-à-vis des protéines, telles que définies ci-
15 dessus.

- La présente invention a pour objet un procédé de dépistage immunologique différentiel de séquences rétrovirales endogènes humaines de la famille HERV-7q normales ou pathologiques, caractérisé en ce qu'il comprend la mise en contact d'un échantillon biologique avec un anticorps selon l'invention, la lecture
20 du résultat étant révélée par un moyen approprié, notamment EIA, ELISA, RIA, fluorescence.

- A titre d'illustration, une telle méthode de diagnostic *in vitro* selon l'invention comprend la mise en contact d'un échantillon biologique prélevé chez un patient, avec des anticorps selon l'invention et la détection à l'aide de tout procédé
25 approprié, notamment à l'aide d'anti-immunoglobulines marquée, des complexes immunologiques formés entre les protéines produites normalement ou pathologiquement et les anticorps.

- Des anticorps monoclonaux ou polyclonaux, produits à partir d'antigènes correspondants à des peptides de synthèse, de polypeptide ou protéines
30 recombinants, permettent de suivre l'expression des peptides ou protéines produits normalement ou pathologiquement. L'analyse est de préférence effectuée par ELISA,

ou équivalent, Western-blot ou équivalent, ou par immunohistochimie.

Les peptides ou protéines, issus des séquences rétrovirales endogènes ou dont l'expression est associée à l'expression de ces séquences rétrovirales endogènes, sont recherchés et identifiés.

5 La présente invention a également pour objet un procédé d'identification et de détection de motifs rétroviraux endogènes, anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier auto-immunes, au premier rang desquelles la sclérose en plaques, caractérisé en ce qu'il comprend l'analyse comparée des séquences extraites d'un échantillon biologique
10 avec les séquences selon l'invention.

La présente invention a également pour objet l'application des séquences nucléiques ou des séquences protéiques selon l'invention au diagnostic, au pronostic, à l'évaluation de la susceptibilité génétique, à toutes maladies humaines induites, innées ou acquises en particulier celles à composantes cancéreuses, auto-
15 immunes et/ou à incidence neurologique, comme la sclérose en plaques, les syndromes associés et les maladies neurodégénératives où intervient tout ou partie des séquences nucléiques selon l'invention et des formes endogènes ou exogènes apparentées.

La présente invention a également pour objet des séquences
20 nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs nucléiques selon l'invention, combinés avec des séquences ou motifs d'origine endogène ou d'origine ou induits de manière exogène.

La présente invention a, en outre, pour objet un vecteur recombinant de clonage ou d'expression, caractérisé en ce qu'il comprend une séquence nucléique
25 conforme à l'invention.

Des manœuvres thérapeutiques peuvent être envisagées par usage de certaines des séquences nucléiques contenues dans HERV-7q et les séquences de la même famille ou des structures polypeptidiques déduites ou par utilisation de peptides ou protéines, ou d'anticorps spécifiques.

30 Conformément à l'invention, tout ou partie des séquences nucléiques rétrovirales endogènes de type HERV-7q, peut être utilisée à usage de vecteur ou

d'éléments de vecteurs à vocation thérapeutique, en particulier les séquences LTR et la région *gag* (SEQ ID NO :2, 21 et 22)..

L'intérêt de telles séquences réside, dans l'innocuité du vecteur ainsi formé, dans la possibilité d'une insertion spécifique ciblée dans une région bien définie par une stratégie analogue à la recombinaison homologue, dans le ciblage cellulaire, éventuellement transitoire dans le cas d'une expression placentaire chez la femme. Un autre aspect concerne la possibilité d'associer aux gènes d'intérêts les motifs rétroviraux biologiquement actifs (peptides immunomodulateurs, tels que représentés aux séquences SEQ ID NO 68-118, ci-après, peptide fusogène...).

La présente invention a également pour objet des animaux transgéniques, caractérisés en ce qu'ils comprennent tout ou partie d'une séquence de type HERV-7q (SEQ ID NO:1-22 et 61).

Le Tableau II ci-après établit les correspondances entre les numéros des séquences telles qu'elles apparaissent dans la liste de séquences et le nom des différentes séquences.

TABLEAU II

SEQ ID NO :	DÉSIGNATION
1	Acide nucléique : 7 env
2	Acide nucléique : gag
3	Acide nucléique : HERV-7q
4	Acide nucléique : HE2
5	Acide nucléique : HE3
6	Acide nucléique : HG3
7	Acide nucléique : HE4
8	Acide nucléique : HE5
9	Acide nucléique : HE6
10	Acide nucléique : HG6
11	Acide nucléique : HE7
12	Acide nucléique : HE8
13	Acide nucléique : HG8
14	Acide nucléique : HE9
15	Acide nucléique : HE10
16	Acide nucléique : HE11
17	Acide nucléique : HG11
18	Acide nucléique : HE12
19	Acide nucléique : HG12

SEQ ID NO:	DÉSIGNATION
20	Acide nucléique : R1
21	Acide nucléique : R1F
22	Acide nucléique + protéine env déduite : HERV-7q
23	Fragment de protéine env déduite selon SEQ ID NO :22
24	Fragment de protéine env déduite selon SEQ ID NO :22
25	Fragment de protéine env déduite selon SEQ ID NO :22
26	Protéine : envérine
27	Fragment de protéine env déduite selon SEQ ID NO :22
28	Acide nucléique + protéine déduite de gag : HERV-7q
29	Fragment de protéine gag déduite selon SEQ ID NO :28
30	Fragment de protéine gag déduite selon SEQ ID NO :28
31	Fragment de protéine gag déduite selon SEQ ID NO :28
32	Fragment de protéine gag déduite selon SEQ ID NO :28
33	Fragment de protéine gag déduite selon SEQ ID NO :28
34	Fragment de protéine gag déduite selon SEQ ID NO :28
35	Protéine env : cadre de lecture 1
36	Protéine gag
37	Acide nucléique : G1F (amorce)
38	Acide nucléique : G1R (amorce)
39	Acide nucléique : G2F (amorce)
40	Acide nucléique : G2R (amorce)
41	Acide nucléique : G4F (amorce)
42	Acide nucléique : G3F (amorce)
43	Acide nucléique : G4R (amorce)
44	Acide nucléique : G5R (amorce)
45	Acide nucléique : E1F (amorce)
46	Acide nucléique : E1R (amorce)
47	Acide nucléique : E2F (amorce)
48	Acide nucléique : E2R (amorce)
49	Acide nucléique : E3F (amorce)
50	Acide nucléique : E3R (amorce)
51	Acide nucléique : E4F (amorce)
52	Acide nucléique : E4R (amorce)
53	Acide nucléique : E5F (amorce)
54	Acide nucléique : E6F (amorce)
55	Acide nucléique : E5R (amorce)
56	Acide nucléique : ExF (amorce)
57	Acide nucléique : ExR (amorce)
58	Protéine gag
59	Acide nucléique : Séquence A (séquence d'insertion)
60	Acide nucléique : Séquence B (séquence d'insertion)
61	Acide nucléique : HE13
62	Acide nucléique : RH7

SEQ ID NO:	DÉSIGNATION
63	Acide nucléique : RAM75
64	Acide nucléique : RAV73
65	Acide nucléique : RBP3
66	Acide nucléique : HI3
67	Acide nucléique : LTX
68	Peptide Tableau I
69	Peptide Tableau I
70	Peptide Tableau I
71	Peptide Tableau I
72	Peptide Tableau I
73	Peptide Tableau I
74	Peptide Tableau I
75	Peptide Tableau I
76	Peptide Tableau I
77	Peptide Tableau I
78	Peptide Tableau I
79	Peptide Tableau I
80	Peptide Tableau I
81	Peptide Tableau I
82	Peptide Tableau I
83	Peptide Tableau I
84	Peptide Tableau I
85	Peptide Tableau I
86	Peptide Tableau I
87	Peptide Tableau I
88	Peptide Tableau I
89	Peptide Tableau I
90	Peptide Tableau I
91	Peptide Tableau I
92	Peptide Tableau I
93	Peptide Tableau I
94	Peptide Tableau I
95	Peptide Tableau I
96	Peptide Tableau I
97	Peptide Tableau I
98	Peptide Tableau I
99	Peptide Tableau I
100	Peptide Tableau I
101	Peptide Tableau I
102	Peptide Tableau I
103	Peptide Tableau I
104	Peptide Tableau I
105	Peptide Tableau I

106	Peptide Tableau I
SEQ ID NO:	DÉSIGNATION
107	Peptide Tableau I
108	Peptide Tableau I
109	Peptide Tableau I
110	Peptide Tableau I
111	Peptide Tableau I
112	Peptide Tableau I
113	Peptide Tableau I
114	Peptide Tableau I
115	Peptide Tableau I
116	Peptide Tableau I
117	Peptide Tableau I
118	Peptide Tableau I
119	Acide nucléique : BLIMP-1
120	Peptide : CKH
121	Acide nucléique : F645 (amorçe)
122	Acide nucléique : PS5D (amorçe)

Outre les dispositions qui précèdent, l'invention comprend encore d'autres dispositions, qui ressortiront de la description qui va suivre, qui se réfère à des exemples de mise en œuvre du procédé objet de la présente invention ainsi qu'aux
5 dessins annexés, dans lesquels :

- Figure 1. Séquence nucléique humaine HERV-7q, dont l'analyse et le traitement permettent de caractériser une nouvelle structure rétrovirale endogène. Les régions nucléiques répétées de type R1 et R2 et les domaines *gag*, *pol* et *env* sont soulignés. Les domaine de type *gag* et *env* sont en italiques. La région homologe à
10 une partie 3' non-codante de Rab7 est doublement soulignée.

- Figure 2. Cartographie de la région rétrovirale endogène humaine HERV-7q. La partie haute de la figure correspond à une région anonyme du génome humain située sur le bras long du chromosome 7. On peut identifier les domaines répétés (1), *gag* (2), *pol* (3) et *env* (4) de HERV-7q. La région *env* C-terminale (4.3) se
15 prolonge en amont en un long cadre de lecture ouvert (4.2). Le domaine 4.1, correspond à la région N-terminale du domaine *env*.

- Figure 3. Comparaison des séquences nucléiques répétées situées aux bornes de HERV-7q. Les régions nucléiques répétées 5'(haut) et 3'(bas), sont

comparées et les bases identiques sont indiquées par deux points.

- Figure 4. Séquence déduite présentant un cadre de lecture ouvert, dans le domaine de type-*env* de HERV-7q selon la règle du plus long cadre de lecture ouvert.

5 - Figure 5. Séquences autour du domaine CKS-17 identifiées dans différents domaines *env* déduits de la famille de HERV-7q et comparaison avec des motifs CKS-17 de référence.

1) HE2 - 2) HERV-7q - 3) N° d'accès à GenBank: M85205 - 4) HE7 - 5) HE9 - 6) CKS-17: le motif peptidique doué de propriétés immunomodula-
10 trices est souligné - 7) gp20 de rétrovirus de type-D (SRV-Pc).

- Figure 6. Séquence déduite possible du domaine de type-*gag* identifié dans HERV-7q établie selon la règle du plus long cadre de lecture ouvert. X et / correspondent respectivement à un codon non-sens et à un décalage de cadre de lecture. La séquence soulignée correspond au début du domaine *pol*.

15 - Figure 7. Comparaison des régions nucléiques couvrant la région *gag* de HERV-7q (haut) et HERV-TcR (bas) et leurs régions flanquantes. Les bases identiques sont spécifiées par deux points.

- Figure 8. Exemple d'alignements nucléiques du domaine de type *env* de HERV-7q avec des domaines de type *env* similaires présents dans des
20 séquences rétrovirales endogènes humaines de la même famille. Les codons non sens sont soulignés : 1) HERV-7q - 2) HE2 - 03) HE3 - 04) HE4.

- Figure 9. Alignements nucléiques entre le domaine *gag* de HERV-7q et les domaines correspondants appartenant à la même famille. Comparaison avec des fragments de domaines *gag* isolés d'agents rétroviraux infectieux. Séquences
25 d'origine rétrovirale infectieuse: N° d'accès dans la banque de données EMBL : 1) A60168 - 2) A60201 - 3) A60200 - 4) A60171. Séquences rétrovirales endogènes humaines: 5) HERV-7q - 6) HG11 - 7) HG3. Les chiffres indiqués dans les séquences endogènes, correspondent au nombre de nucléotides insérés afin d'optimiser l'alignement avec les séquences de type *gag* identifiées dans des rétrovirus d'origine
30 infectieuse.

- Figure 10. Alignement d'un motif *gag* protéique déduit (haut)

appartenant à un rétrovirus infectieux (N° d'accension EMBL : A60200) avec le motif *gag* protéique déduit (bas) identifié dans HERV-7q. Les codons non-sens sont en gras et soulignés. Les acides aminés identiques sont spécifiés par 2 tirets. Un tiret indique une délétion ou un acide aminé homologue.

5 - Figure 11. Alignement d'un motif *env* (haut) appartenant à un rétrovirus infectieux (N° d'accension EMBL : A60170) avec le motif *env* (bas) identifié dans HERV-7q. Les nucléotides homologues sont spécifiés par deux points et les délétions par un tiret.

 - Figure 12. Comparaison entre le domaine *env* de HERV-7q (haut)
10 et le domaine *env* de HERV-9 (bas). L'homologie de 66 % se limite à la région 3' du domaine *env* de HERV-7q et HERV-9, respectivement entre les nucléotides 8976 nt et 9500 nt de HERV-7q et les nucléotides 2898 nt et 3465 nt de HERV-9 (N° d'accension à GenBank : X57147). De nombreuses insertions/délétions sont aussi observées.

15 - Figure 13. Homologie entre une partie de la séquence du transcrit codant pour RH7 (haut, SEQ ID NO:62) et un motif de RGH2 (bas - N° d'accension à GenBank: D11018).

 - Figure 14. Identification de la séquence du transcrit codant pour RAM75 (SEQ ID NO:63), correspondant au gène d'une ATPase de type PEX1. Les
20 exons codants sont soulignés. Les codons d'initiation et non-sens ainsi que les sites putatifs de polyadénylation sont en gras et soulignés. La région en italique correspond au début de la séquence rétrovirale endogène RH7.

 - Figure 15. Séquence du transcrit codant pour RAV73 (SEQ ID NO:64), située à 0.7 kb en aval de HERV-7q ; les séquences nucléiques aptes à coder
25 pour un ou plusieurs polypeptides sont soulignées.

 - Figure 16. Comparaison entre la séquence 3' LTR (haut) de HERV-7q et la séquence intronique LTX (SEQ ID NO:67), située dans le gène FMR2, associé au X-fragile (bas).

 - Figure 17. Mise en évidence de modifications sur la séquence
30 nucléotidique (ID NO:3), chez des patients atteints de SEP. Les bases modifiées, chez au moins un patient, sont soulignées. Les amorces utilisées sont en italiques

(séquences SEQ ID NO:121 et 122). L'ATG d'initiation et le codon non-sens sont en gras.

- Figure 18. Partie codante *env* de la séquence HERV-7q (séquence ID NO:3), avec 3 cadres de lecture.

5 - Figures 19, 20, 21. Présentation séparée de la protéine *env* selon les 3 cadres de lecture.

- Figure 22. Séquence nucléique contenant la séquence rétrovirale RH7 située en 5' de la séquence HERV-7q. La séquence en italique correspond au début de la séquence HERV-7q. La séquence RH7 est soulignée. Deux sites de poly-
10 adénylation putatifs sont gras.

- Figure 23. Séquence du transcrit codant pour RBP3 contenant des motifs nucléotidiques identifiés dans la séquence nucléique codant pour le gène Blimp-1.

- Figure 24. Séquence du transcrit codant pour APS.

15 - Figure 25. Séquence du transcrit codant pour Blimp-1 ; la partie codante est soulignée ; les codons d'initiation et de terminaison sont en gras.

- Figure 26. Séquence du transcrit codant pour FMR2. La partie codante est soulignée. Les codons d'initiation et non-sens sont en gras.

Il doit être bien entendu, toutefois, que ces exemples sont donnés
20 uniquement à titre d'illustration de l'objet de l'invention, dont ils ne constituent en aucune manière une limitation.

EXEMPLE 1 : Détection, par amplification génique, d'une séquence nucléique appartenant à un domaine de type *gag* ou *env* selon l'invention, dans un échantillon d'ADN génomique d'origine humaine ou de mammifères.

25 L'amplification génique s'effectue à partir d'ADN génomique isolé à partir du sang. Un traitement anticoagulant est effectué avec 1 ml d'une solution de citrate (pour un litre : 4,8 g de d'acide citrique, 13,2 g de citrate de sodium, 14,7 g de glucose) pour 6 ml de sang frais. Après centrifugation de 20 ml de sang pendant 15 mn à 13.0000 g, le surnageant est éliminé et la fraction enrichie en globules blancs est

transférée dans un nouveau tube, puis recentrifugée dans les mêmes conditions que précédemment. La fraction enrichie en globules blancs est resuspendue dans un tampon d'extraction (10 mM Tris-HCl, 0,1 M EDTA, 20 µg/ml de RNase pancréatique traitée afin d'éliminer les DNases, 0,5 % SDS, pH 8,0), puis incubée pendant 1 heure à 37°C. La protéinase K est ajoutée à une concentration finale de 100 µg/ml. La suspension des cellules lysées est incubée à 50°C durant 3 heures sous agitation périodique, puis traitée par un volume égal de phénol équilibré par du Tris-HCl 0,5 M, pH 8,0. L'émulsion formée est placée sur une roue pendant une heure, puis centrifugée à 5000 g pendant 15 mn à température ambiante. La solution aqueuse est traitée déprotéinisée par une triple extraction phénolique afin d'obtenir un niveau de purification correspondant à un rapport final d'absorbance A260/A280 supérieur à 1,75. La fraction aqueuse est précipitée par 0,2 vol. d'acétate de sodium 10 M et 2 vol. d'éthanol. L'ADN est alors soit prélevé avec l'extrémité d'une pipette pasteur recourbée, soit centrifugé à 5000 g pendant 5 mn à température ambiante. L'ADN ou le culot d'ADN est lavé deux fois par de l'éthanol à 70 %, puis repris dans 1 ml de TE pH 8,0 afin d'être élué sous agitation douce pendant 12 à 24 heures.

Des oligonucléotides spécifiques des séquences endogènes décrites selon l'invention sont choisis pour amplifier la région *gag* ou *env* des régions rétrovirales endogènes décrites selon l'invention. L'ADN génomique étudié provient de patients présentant des pathologies comme la sclérose en plaques et d'individus réputés sains.

Les ADN polymérases thermostables utilisées ont été choisies pour leur grande fidélité lors du processus d'amplification, comme la Vent, ADN polymérase (Biolabs) ou équivalent, et sont utilisées selon les conditions préconisées par le fournisseur.

La stratégie d'amplification utilise selon les cas une simple PCR, ou une PCR nichée ou semi-nichée.

Oligonucléotides utilisés pour amplifier la région *gag* :

- amorce G1F, sens, localisée dans la région amont du domaine *gag* de *HERV-7q* (SEQ ID NO:37),

- amorce G1R, anti-sens, localisée dans la région 3' terminale du

domaine *gag* (SEQ ID NO:38),

Le fragment de 1505 nt amplifié par le couple G1F-G1R : 1505 nt est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR.

- 5 - amorce G2F, sens nichée (SEQ ID NO:39),
 - amorce G2R, anti-sens nichée (SEQ ID NO:40),
 - amorce G4F, sens nichée (SEQ ID NO:41),
 - amorce G3F, sens nichée (SEQ ID NO:42),
 - amorce G4R, anti-sens nichée (SEQ ID NO:43),
10 - amorce G5R, anti-sens nichée (SEQ ID NO:44),

Oligonucléotides utilisés pour amplifier la région *env* de HERV-7q :

- amorce E1F, sens (SEQ ID NO:45),
 - amorce E1R, anti-sens (SEQ ID NO:46),
Le fragment de 2529 nt amplifié par le couple d'amorces E1F-E1R,
15 est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR.

- amorce E2F, sens (SEQ ID NO:47),
 - amorce E2R, antisens (SEQ ID NO:48),
 - amorce E3F, sens (SEQ ID NO:49),
20 - amorce E3R, anti-sens (SEQ ID NO:50),
 - amorce E4F, sens (SEQ ID NO:51),
 - amorce E4R, anti-sens (SEQ ID NO:52),
 - amorce E5F, sens (SEQ ID NO:53),
 - amorce E6F, sens (SEQ ID NO:54)
25 - amorce E5R (SEQ ID NO:55).
 - amorce ExF (SEQ ID NO:56)
 - amorce ExR (SEQ ID NO:57)

La PCR est réalisée à partir de 50 à 200 ng d'ADN génomique. Les conditions de PCR sont celles préconisées par le fournisseur. Les conditions cycliques
30 d'amplification sont réalisées dans 50 µl : une dénaturation de 94°C pendant 1 min.,
une hybridation de 70°C pendant 1 min., et une élongation à 72 °C pendant 1 à 2 min.,

selon les fragments amplifiés. Après 35 cycles, une réaction terminale est menée à 72°C pendant 10 min. Le séquençage automatique des échantillons amplifiés est réalisé à l'aide d'un séquenceur Applied Biosystems de type ABI 377 ou autre modèle comparable, selon les protocoles fournis par le constructeur.

- 5 Dans le cas d'une PCR nichée ou semi-nichée, les mêmes conditions expérimentales sont utilisées, à la seule différence que l'échantillon d'ADN génomique est remplacé par 5 à 10 µl du produit d'amplification issu de la première PCR.

Deux amplifications indépendantes sont réalisées à partir du même échantillon. Une réaction de contrôle est réalisée en remplaçant l'échantillon d'ADN
10 par de l'eau afin de détecter d'éventuels contaminants.

EXEMPLE 2 : Détection par amplification génique d'une séquence nucléique selon l'invention dans un échantillon biologique d'ADN génomique prélevé chez des patients présentant une pathologie candidate déclarée ou la suspicion de cette pathologie.

- 15 Le protocole d'amplification est le même que dans l'exemple 1, mis à part l'origine de l'échantillon qui provient de patients présentant une pathologie candidate. Un échantillon d'ADN génomique réputé normal est systématiquement intégré dans l'ensemble des échantillons pathologiques amplifiés puis analysés.

Les produits de PCR sont séparés sur un gel d'agarose à 1,5 %, puis
20 transférés en présence de soude 0,4 N sur une membrane de nylon chargé. Une hybridation est réalisée avec une sonde spécifique correspondant aux fragments de PCR amplifiés soit par les couples G1F-G1R soit par le couple E1F-E1R. La sonde est marquée par incorporation de dUTP-digoxygénine selon le protocole du fournisseur (Boehringer Mannheim). L'hybridation est effectuée dans un tampon d'hybridation
25 (5XSSC, 50 % formamide, 0,1 % lauroyl-sarcosine, 0,02 % SDS, 2 % de réactif de blocage Boehringer) pendant une nuit à 42°C. Le Southern est lavé 2 fois 5 min. à température ambiante dans une solution de 2XSSC, 0,1% SDS. Puis un lavage à haute stringence est effectué à deux reprises pendant 15 min. à 55°C dans une solution 0,1XSSC, 0,1 % SDS. L'hybridation est révélée selon le protocole du fournisseur
30 (Boehringer Mannheim), en présence d'un substrat chimioluminescent de la phosphatase alcaline, de type CSPD ou CDP-STAR. Le filtre est révélé après une exposition

de 15min. à 60 min.

Une analyse par SSCP ("*single strand conformation polymorphism*") permet de détecter des modifications discrètes de la séquence des fragments amplifiés par PCR. La PCR est menée en présence de dCTP marqués au P³². L'échantillon à analyser est dénaturé à 95°C pendant 10 min., en présence de tampon de charge, puis immédiatement chargé sur un gel de polyacrylamide à 10%, contenant 7.5% de glycérol. La migration s'effectue à 4°C à 8-10 W. Le gel est séché puis autoradiographié.

Les fragments de PCR susceptibles de présenter une altération de leur séquence nucléotidique sont séquencés selon l'exemple 1.

Une hybridation à l'aide d'un oligonucléotide spécifique (17 mers à 20 mers) correspondant à la région nucléotidique modifiée permet d'identifier les échantillons présentant une modification identique (méthode ASO). Brièvement le southern est hybridé avec un oligonucléotide marqué distalement soit au P³², soit en présence de digoxygénine (selon le protocole de Boehringer Mannheim) puis lavé dans des conditions stringentes à 65°C dans une solution 6XSSC, 0.05% pyrophosphate de sodium.

Par exemple, un séquençage nucléotidique automatique a été réalisé sur six fragments de PCR, provenant de 5 patients atteints de SEP et un témoin réputé normal, et qui ont été amplifiés à partir des amorces F645: CTTCAAACAACAACCAGGAGG (SEQ ID NO:121) (située à 26 nucléotides en amont de la méthionine d'initiation de l'envérine) et PS5D: TTGGGGAGGTTGGCCGACGA (SEQ ID NO:122) (située à 6 nucléotides en aval du codon non-sens de l'envérine. Des modifications de la séquence de l'envérine ont été observés sur l'ADN de certains des patients (figure 17).

EXEMPLE 3 : Détection d'une protéine selon l'invention dans un échantillon biologique.

- Préparation d'une fraction protéique purifiée de liquide céphalo-rachidien de patients atteints de SEP

Après un traitement à 56°C pendant 30 min, et élimination des immunoglobulines sur une colonne de protéine G HiTrap (Pharmacia), l'équivalent de

10 ml de LCR est déposé sur une colonne de DEAE Sepharose CL-6B (Pharmacia). L'élution est réalisée en Tris-HCl 20 mM pH 8,8, et un gradient de 0 à 0,4 M de NaCl, puis la fraction est dialysée 2 fois contre du tampon phosphate-NaCl (PBS). Après concentration sur Ultrafree-MC (Millipore), la fraction est déposée sur une colonne de
5 Superose 12 (FPLC Pharmacia) et éluée en présence de PBS. Après séparation par électrophorèse en gel de polyacrylamide-SDS, et électro-transfert sur une membrane d'Immobilon-P (Millipore), les bandes protéiques sont soumises à une hydrolyse trypsique ménagée.

- Analyse de la fraction protéique par spectrométrie de masse

10 Les peptides digérés en présence de trypsine, sont analysés par la méthode de MALDI-TOF, qui permet l'analyse de peptides présents en mélange. (COTTRELL J.S., Pept. Res., 1997, 7, 115-124). Les peptides caractérisés en fonction de leur masse sont comparés aux protéines et aux protéines associées selon l'invention.

15 **EXEMPLE 4 : Détection d'anticorps spécifiques anti-domaine *env* de HERV-7q.**

L'identification d'un long cadre de lecture ouvert au sein de la séquence *env* de HERV-7q, a permis de déterminer une séquence protéique déduite SEQ ID NO:22 et 35 et figures 18-20 d'une région dudit gène.

20 Les séquences de protéines déduites des séquences ID NO:22, 35 et des figures 18-20 sont positionnées comme suit par rapport à la figure 1 ou à la séquence ID NO:3 :

SEQ ID NO:22 (cadre de lecture 1) et figure 19 : début de la séquence codante : position 7874, fin de la séquence codante 1^{er} codon non-sens (position 9493)

25 SEQ ID NO:35 : début de la séquence codante : position 7874, fin de la séquence codante 1^{er} codon non-sens (position 9493) (cadre de lecture 1)

Figure 19 : début de la séquence codante : position 6970, fin de la séquence codante 1^{er} codon non-sens (position 9493) (cadre de lecture 1)

30 Figure 20 : début de la séquence codante : position 6971, la fin du cadre de lecture est décalée selon le cas de 1, 2 ou 3 codons

Figure 21 : début de la séquence codante : position 6972, la fin du

cadre de lecture est décalée selon le cas de 1, 2 ou 3 codons

Différents peptides correspondant à tout ou partie des SEQ ID NO:22 (voir SEQ ID NO:23-27 et 35) ont été synthétisés par génie génétique afin de tester leur spécificité antigénique vis-à-vis de séra ou de tissus de patients atteints de SEP, par exemple. Brièvement, tout ou partie de la région env de HERV-7q est sous clonée dans les vecteurs pQE30, 31 et 32. Les vecteurs pQE30, 31 et 32 contiennent en 5' du multi-site de clonage les séquences consensuelles pour la transcription (le promoteur fort du bactériophage T5, 2 opérateurs de l'opéron lactose), la traduction (un site d'accrochage ribosomal synthétique). De même, pQE30, 31 et 32 possèdent en 3', le terminateur de transcription du phage λ ainsi qu'un codon "Stop" pour la traduction. L'expression de la protéine s'effectue après transformation dans *E. coli* M15. Le plasmide pQE30, 31 et 32 possèdent en amont du site de polyclonage la séquence codante pour une suite de 6 histidines présentant une affinité pour les ions nickel. Cet enchaînement permet la purification de la protéine chimérique exprimée, par adsorption sur une résine constituée d'un ligand chélatant, l'acide nitrilotriacétique (NTA), chargé de 4 ions nickel (résine NI-NTA, Qiagen).

La transformation s'effectue par électroporation ou traitement au chlorure de calcium. Par exemple, une colonie d'*E. coli* M15 est incubée dans 100 ml de milieu LB contenant 250 μ g de kanamycine, sous agitation à 37°C jusqu'à l'obtention d'une DO^{600} de 0,5. Après une centrifugation de 5 minutes à 2000g à 4°C, le culot bactérien est repris dans 30 ml de solution TFB1 (100 mM de chlorure de rubidium, 50 mM de chlorure de manganèse, 30 mM d'acétate de potassium, 10 mM $CaCl_2$, 15% glycérol, pH 5.8), à 4°C pendant 90 minutes. Après une centrifugation de 5 minutes à 2000g à 4°C, le culot bactérien est repris dans 4 ml de solution TFB2 (10 mM de chlorure de rubidium, 10 mM de MOPS, 75 mM $CaCl_2$, 15% de glycérol, pH 8). Les cellules peuvent être gardées à -70°C par aliquot de 500 μ l. 20 μ l de la ligation et 125 μ l de cellules compétentes sont mélangés et placés dans la glace 20 minutes. Après un choc thermique de 42°C pendant 90 secondes, les cellules sont agitées 90 minutes à 37°C dans 500 ml de milieu Psi-broth (milieu LB complémenté par 4 mM de $MgSO_4$, 10mM de chlorure de potassium). Les cellules transformées sont étalées sur des boîtes LB-agar complémentées par 25 μ g/ml de kanamycine, et 100 μ g/ml

d'ampicilline, et les boîtes sont incubées une nuit à 37°C.

Les clones potentiellement recombinants sont repiqués de manière ordonnée sur un filtre de nylon déposé sur une boîte LB-agar complémentée par 25 µg/ml de kanamycine et 100 µg/ml d'ampicilline. Après une nuit à 37°C, les clones
5 recombinants sont repérés par hybridation de l'ADN plasmidique avec la sonde nucléotidique amplifiée par PCR avec le couple d'amorces selon SEQ ID NO:45 et SEQ ID NO:46.

Une colonie indépendante, contenant l'insert, est inoculée à 20 ml de milieu LB complémentée par 25 µg/ml de kanamycine et 100 µg/ml d'ampicilline.
10 Après une nuit à 37°C sous agitation, 500 ml de même milieu sont incubés au 1/50° par cette préculture jusqu'à l'obtention d'une D0⁶⁰⁰ de 0,8, puis 1 à 2 mM final d'IPTG est ajouté. Après 5 heures, les cellules sont centrifugées 20 minutes à 4000 g.

Une partie du culot cellulaire est repris dans 5 ml de tampon de sonication (50 mM de phosphate de sodium pH 7,8, 300 mM NaCl) puis placé dans la
15 glace. Après une rapide sonication, les cellules sont centrifugées 20 minutes à 10000 g. Une partie du culot cellulaire est repris dans 10 ml d'une solution 30 mM Tris/HCl-20% sucrose pH8. Les cellules sont incubées 5 à 10 minutes sous agitation, après adjonction de 1 mM EDTA. Après une centrifugation de 10 minutes à 8000 g à 4°C, le culot est repris dans 10 ml de 5 mM de MgSO4 glacé. Après 10 minutes dans
20 la glace sous agitation, les cellules sont centrifugées 10 minutes à 8000 g à 4°C.

Le culot est repris par 5 ml/g dans du tampon A (6 M GuHCl (chlorhydrate de guanidine), 0,1M phosphate de sodium, 0,01M Tris/HCl, pH 8), 1 heure à température ambiante. Le lysat est centrifugé 15 minutes à 10000 g à 4°C, et le surnageant est complémenté par 8 ml de résine Ni-NTA, prééquilibrée dans du tampon
25 A. Après 45 minutes à température ambiante, la résine est coulée dans une colonne, lavée par 10 fois le volume de la colonne par du tampon A puis par 5 fois le volume de la colonne par du tampon B (8 M urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 8). La colonne est lavé par du tampon C (8 M urée, 0,1M phosphate de sodium, 0,01 M Tris/HCl, pH 6,3) jusqu'à ce que l'A280 soit inférieur à 0,01. La
30 protéine recombinante est éluée par 10 à 20 ml de tampon D (8 M urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 5,9) puis par 10 à 20 ml de tampon E (8 M

urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 4,5), puis par 20 ml de tampon F (6 M HCl, 0,2 M acide acétique). Après une analyse en SDS-PAGE, la ou les fractions purifiées contenant la protéine chimérique ont permis l'obtention d'anticorps chez le lapin. Les anticorps obtenus sont testés par Western-blot après révélation par un anticorps secondaire couplé à la phosphatase alcaline.

Des anticorps sont obtenus de la même manière, à partir de peptides synthétisés chimiquement selon la technique de Merrifield (G. Barany and B. Merrifield, 1980, dans *The peptides*, 2, 1-284, E. Gross et J. Meienhofer, Academic Press, New York).

Les anticorps spécifiques obtenus sont utilisés à fin de détection de l'expression sérique ou tissulaire de tout ou partie des séquences rétrovirales endogènes selon l'invention, dans les cas normaux et pathologiques.

Les protéines d'origine sérique ou tissulaire, sont séparées sur gel d'acrylamide-SDS puis transférées sur un filtre de nitrocellulose à l'aide d'un appareil Novablot 2117-2250 (LKB). Le transfert est effectué sur une feuille de Hybond C-extra (Amersham) en utilisant un tampon CAPS 100 mM pH 11, méthanol, eau (V/V/V: 1/1/8) contenant 1 mM de CaCl_2 . Après un transfert de 1 heure à 0,8 mA/cm², la feuille est saturée une heure à température ambiante dans du PBS-0,5 % gélatine. La feuille est mise en présence de l'anticorps spécifique à la concentration de 1/1000 dans du PBS-0,25 % gélatine. Au bout de 2 heures, le filtre est lavé 3 fois 15 minutes dans du PBS-0,1 % de Tween-20, puis le filtre est incubé 30 minutes en présence d'un anticorps secondaire couplé à la phosphatase alcaline (Promega), dilué au 1/7500 dans du PBS-0,25% gélatine. Après trois lavages dans du PBS-0,1 % de Tween-20, le filtre est équilibré dans un tampon (100 mM de Tris-HCl pH 9,5, 100 mM de NaCl, 5 mM de MgCl_2). La révélation est effectuée en présence de 45 µl de NBT à 75 mg/ml et 35 µl de BCIP à 50 mg/ml, pour 10 ml de tampon de phosphatase alcaline.

Les protéines chimériques obtenues par génie génétique, sont utilisées aussi à fin de tests d'activité biologique, comme par exemple pour le test d'activité biologique du peptide de type CKS-17 identifié dans le domaine *env* de

HERV-7q (figure 5).

EXEMPLE 5 : Obtention de sondes ribonucléiques codant pour les séquences env de HERV-7q.

Les fragments de PCR obtenus sont sous clonés dans le plasmide
5 PGEM 4Z (Promega) qui possède de par et d'autre de son site de polyclonage, les séquences promotrices pour les ARN polymérase SP6 et T7.

La méthode de compétence utilisée est l'électroporation. Le plasmide et le fragment de PCR sont hybridés dans un rapport de 50 ng de vecteur (coupé à Sma I) pour 100 ng de fragment de PCR (rendu à bout franc par traitement par le fragment
10 de Klenow de l'ADN polymérase). L'incubation a lieu une nuit à 22°C, dans le tampon de ligation (66 mM Tris-HCl pH 7,5, 5 mM MgCl₂, 1 mM dithioerythritol, 1 mM ATP) en présence de 1u. de T4 ADN ligase puis est arrêtée par dénaturation 10 minutes à 65°C. Parallèlement, la souche d'*E. Coli* JM 105 estensemencée une nuit à 37°C dans du milieu LB. Cette préculture est diluée au 1/500 et placée à 37°C jusqu'à
15 une DO⁶⁰⁰ égale à 1. Pour la suite du mode opératoire les cellules seront toujours conservées au froid. Après une centrifugation de 5 minutes à 3500 g à 4°C, le culot cellulaire est resuspendu dans 1/4 vol. d'eau glacée ultra-pure. Cette étape est répétée 5 à 6 fois. Puis le culot est resuspendu dans 1/4000 vol. d'eau; 10 % de glycérol stérile sont ajoutés permettant la conservation des cellules électrocompétentes, par aliquots
20 de 10 µl à 20°C. A 50 µl de cellules électrocompétentes est ajouté 1 µl de la ligation ; le tout est soumis à une décharge électrique de 12,5 kV/cm, appliquée pendant 5,8 ms. Les cellules sont rapidement remises en suspension dans le milieu SOC, incubées 1 heure à 37°C, puis étalées, en présence de 2% X-Gal dans du diméthylformamide, et 10 mM d'IPTG, sur une boîte de gélose LB-agar supplémentée en ampicilline (100
25 µg/ml). Après une nuit à 37°C, les clones blancs potentiellement recombinants, sont repiqués de manière ordonnée sur une boîte LB/ampicilline et parallèlement sur un filtre de nylon déposé sur une boîte LB/ampicilline. Ces deux boîtes sont incubées une nuit à 37°C. Les clones recombinants sont alors repérés par hybridation avec une sonde nucléique amplifiée par PCR avec le couple d'amorces selon SEQ ID NO:45 et
30 SEQ ID NO:46 et marquée à la digoxygénine.

Les clones recombinants sont cultivés dans 50 ml de milieu

LB/ampicilline (100 µg/ml) en agitation pendant une nuit à 37°C. Après une centrifugation à 3500 g pendant 15 minutes à 4°C, le culot bactérien est repris dans 4ml de tampon P1 (50 mM Tris-HCl, 10mM EDTA, 400 µg/ml RNase A, pH 8) et 4ml de tampon P2 (200 mM NaOH, 1% SDS). Le mélange est incubé à température ambiante pendant 5 minutes. Après adjonction de 4ml de tampon P3 (2.55 M d'acétate de potassium, pH 4,8) le mélange est centrifugé à 12000 g pendant 30 minutes à 4°C. Le surnageant est appliqué sur une colonne Qiagen-type 100, prééquilibrée avec 2 ml de tampon QBT (750 mM NaCl, 50 mM MOPS, 15% éthanol, pH 7). la colonne est lavée avec 2 fois 4ml de tampon QC (1M NaCl, 50 mM MOPS, 15 % éthanol, pH 7) et l'ADN est élué avec 2ml de tampon QF (1,2 M NaCl, 50mM MOPS, 15 % éthanol, pH 8). L'ADN est précipité avec 0,8 vol. d'isopropanol, et centrifugé à 12000 g à 4°C pendant 30 minutes. Le culot est lavé avec de l'éthanol à 70 % glacé. puis l'ADN plasmidique est repris par 2 fois 150 µl de tampon TE.

Les sondes ribonucléiques sont utilisées comme sondes spécifiques, en particulier pour la détection des transcrits exprimés par les séquences rétrovirales endogènes selon l'invention.

EXEMPLE 6 : Construction d'une souris transgénique contenant tout ou partie du gène de l'envérine.

Une souris transgénique contenant tout ou partie de la séquence HERV-7q (SEQ ID NO:3) est construite afin d'identifier les séquences responsables de la spécificité tissulaire, et pour évaluer le rôle de tout ou partie des motifs rétroviraux endogènes de type HERV-7q, en particulier tout ou partie des motifs peptidiques de l'envérine. La technique de micro-injection utilisée se réfère à la technique classique (Hogan et coll., (1994), Manipulating the mouse embryo, Cold Spring Harbor, Cold Spring Harbor Laboratory Press) ou à ses équivalents. Des formes identiques à la molécule humaine normale de motifs de type HERV-7q, dont l'envérine, ou des formes mutées, délétées, présentant des insertions ou tronquées sont testées afin de déterminer les motifs critiques tant sur le plan normal que pathologique, et plus particulièrement au cours du développement foetal et lors des processus tumoraux.

30 Bibliographie :

- Benit L. et al., 1997. Cloning of a new murine endogenous retrovirus MuERV-L, with

- strong similarity of the human HERV-L element and with a *gag* coding sequence closely related to the Fv1 restriction gene. J. Virol. 71, 5652-5657.
- Coffin J.M. 1985. Endogenous retrovirus. In: "RNA tumor viruses" (Weiss R.A., Varmus H.E., Teich N.M., and Coffin J.M. eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Conrad B., Weissmahr R.N., Boni J., Arcari R., Schupbach J., and Mach B. 1997. A human endogenous retroviral superantigen as candidate autoimmunogene in type 1 diabetes. Cell 90, 303-313.
- Covey S.N. 1986. Amino acid sequence homology in *gag* region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. Nucleic Acids Res. 14, 623-633.
- Hertig C., Coupar B.E., Gould A.R., and Boyle D.B. 1997. Field and vaccine strains of fowlpox virus carry integrated sequences from the avian retrovirus, reticuloendotheliosis virus. Virology 235, 367-376.
- Hohenadl C., Leib-Mösch C., Hehlemann R., and Erfle Y. 1996. Biological significance of human endogenous retroviral sequences. J. Acqui. Imm. Def. Synd. Hum. Retrovir. 13, S268-S273.
- Kulkoski J.K., Jones S., Katz R.A., Mack J.P.G., and Skalka A.M. 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. 12, 2331-2338.
- La Mantia G. et al, N.A.R., 1991, 19, 7, 1513-1520
- Patience C., Wilkinson D.A., and Weiss R.A. 1997. Our retroviral heritage. Trends Genet. 13, 116-120.
- Pearson W.R. 1994. Using the FASTA program to search protein and DNA sequence databases. Methods Mol. Biol. 24, 307-331.
- Perron H., Garson J.A., Bedin F., Beseme F., Paranhos-Baccala G., Komurian-Pradel F., Mallet F., Tuke P.W., Voisset C., Blond J.L., Lalande B., Seigneurin J.M., Mandrand B. and the Collaborative Research Group on Multiple Sclerosis. 1997. Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. Proc. Natl. Acad. Sci. USA 94, 7583-7588.

- Tönjes R.R. et al., J. AIDS and Hum. Retrovirol, 1996, 13, S261-S267.
- Vitelli R., Chiarillo M., Lattero D., Bruni C.B., and Bucci C. 1996. Molecular cloning and expression analysis of the human Rab7 GTP-ase complementary deoxyribonucleic acid. Biochem. Biophys. Res. Commun. 229, 887-890.
- 5 - Weber L.T., Miller M., Jaskolski M., Leis J., Skalka M., and Wlodawer A. 1989. Molecular modeling of the HIV-1 protease and its substrate binding site. Science 243, 928-931.
- Wilkinson D., Mager D.L., and Leong J.A.C. 1994. Endogenous human retroviruses. In: "The Retroviridae" (Levy, J.A. ed), Plenum Press New York, , Vol. 3, 465-535.
- 10 - Xiong Y., and Eickbush, T. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9, 3353-3362.

Ainsi que cela ressort de ce qui précède, l'invention ne se limite nullement à ceux de ses modes de mise en œuvre, de réalisation et d'application qui viennent d'être décrits de façon plus explicite ; elle en embrasse au contraire toutes les

15 variantes qui peuvent venir à l'esprit du technicien en la matière, sans s'écarter du cadre, ni de la portée, de la présente invention.

REVENDICATIONS

1°) Fragment d'acide nucléique purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70% sur plus de 600 nucléotides pour les domaines de type *env*.

2°) Fragment d'acide nucléique selon la revendication 1, caractérisé en ce qu'il présente à la fois de motifs rétroviraux correspondant à un domaine *env* et répondant à la séquence SEQ ID NO:1 et des motifs rétroviraux correspondant à un domaine *gag* et répondant à la séquence SEQ ID NO:2 ou à une séquence présentant un niveau d'homologie supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70% sur plus de 600 nucléotides pour les domaines de type *env* et un niveau d'homologie supérieur ou égal à 90% sur plus de 700 nucléotides ou supérieur ou égal à 70% sur plus de 1200 nucléotides pour les domaines de type *gag*, lesquels motifs ne présentent aucune insertion ou délétion supérieure à 200 nucléotides.

3°) Fragment d'acide nucléique, caractérisé en ce qu'il comprend un segment d'une séquence selon la revendication 1 ou la revendication 2 et notamment les séquences SEQ ID NO:3-22, 28 et 61, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes ainsi que les fragments issus des régions codantes des séquences précédentes correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

4°) Transcrits, caractérisés en ce qu'ils sont générés à partir des séquences selon l'une quelconque des revendications 1 à 3.

5°) Réactif de diagnostic pour la détection différentielle de séquences nucléiques endogènes humaines complètes ou partielles, présentant des motifs rétroviraux, sélectionnés parmi les séquences SEQ ID NO:1 et/ou SEQ ID NO:2, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1-22, 28, 37-57, 59-61 et 121-122, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes, par les fragments nucléotidiques capables de définir ou d'identifier les séquences SEQ ID

NO:1 et/ou SEQ ID NO:2 et toute séquence flanquante ou les chevauchant ainsi que par les fragments issus des régions codantes des séquences SEQ ID NO:1-22 et 61, correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, éventuellement marquées avec un marqueur approprié.

5 6°) Réactif selon la revendication 5, caractérisé en ce qu'il est choisi dans les régions situées entre les nucléotides 3065 et 4390, les nucléotides 6965 et 9550 ou les nucléotides 2502-2865 de la SEQ ID NO:3.

7°) Réactif selon la revendication 5, caractérisé en ce qu'il est sélectionné parmi les séquences SEQ ID NO:37-57, 59-60 et 121-122 et en ce qu'il est apte
10 à être utilisé comme amorce.

8°) Réactif selon la revendication 5, caractérisé en ce qu'il est sélectionné parmi les séquences suivantes :

- un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

15 - un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R)

- un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865

et en ce qu'il est apte à être utilisé comme sonde.

20 9°) Réactif selon la revendication 5, caractérisé en ce qu'il est choisi dans le groupe constitué par les fragments codants ou non-codants pour tout ou partie de l'envérine et notamment les fragments comprenant au moins 14 nucléotides et plus particulièrement les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter du codon
25 codant pour la première méthionine.

10°) Procédé de détection rapide et différentiel des séquences nucléiques rétrovirales endogènes de type *env* ou *env* et *gag*, de leurs variants normaux ou pathologiques, par hybridation et/ou amplification génique, réalisé à partir d'un échantillon biologique, lequel procédé est caractérisé en ce qu'il comprend :

30 (a) une étape dans laquelle l'on met en contact un échantillon biologique à analyser avec au moins une sonde selon la revendication 5, la revendication 6

ou la revendication 8 et

(b) une étape dans laquelle on détecte par tout moyen approprié le ou les produits résultants de l'interaction séquence nucléotidique-sonde.

11°) Procédé de détection selon la revendication 10, caractérisé en ce qu'il comprend :

* préalablement à l'étape (a) :

. une étape de préparation du tissu ou du liquide biologique concerné,

. une étape d'extraction de l'acide nucléique à détecter, et

10 . au moins un cycle d'amplification génique mis en œuvre à l'aide d'au moins un réactif selon l'une quelconque des revendications 5 à 7 et

* postérieurement à l'étape (b) :

. une étape de comparaison des séquences nucléiques obtenues dans ledit échantillon biologique avec les séquences rétrovirales endogènes humaines selon
15 l'une quelconque des revendications 1 à 3, par tout moyen approprié et notamment par séquençage, Southern-blot, coupure de restriction, SSCP ou toute autre méthode permettant d'identifier une insertion ou une délétion ou encore une simple mutation entre les différentes séquences comparées.

12°) Procédé de détection des transcrits selon la revendication 4, caractérisé en ce qu'il comprend :

- le prélèvement des ARN messagers provenant d'échantillons biologiques témoins et d'échantillons analogues prélevés chez des patients et

- l'analyse qualitative et/ou quantitative desdits ARNm, par hybridation *in situ*, par dot-blot, Northern-blot, RNase mapping ou RT-PCR, à l'aide d'un
25 réactif de diagnostic selon l'une quelconque des revendications 5 à 9.

13°) Séquences chimères, caractérisées en ce qu'elles sont constituées par un fragment de 17 à 40 nucléotides d'une séquence flanquante sélectionnée dans le groupe constitué par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou
30 l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences

correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb, associée à un motif rétroviral endogène de type HERV-7q comprenant entre
5 17 et 40 nucléotides selon les revendications 1 à 4.

14°) Méthode de détection et/ou d'évaluation d'une sur-expression/sous-expression ou d'une modification d'au moins l'une des séquences ou fragments de séquences rétrovirales endogènes de type HERV-7q et/ou de leurs séquences flanquantes associées, selon l'une quelconque des revendications 1 à 9, caractérisée en
10 ce qu'elle comprend :

- le dépôt sur un support approprié, de l'ADNc provenant de clones, de produits de PCR obtenus à partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, lesdites séquences d'ADN étant des séquences ou des fragments de séquences rétro-
15 virales endogènes de type HERV-7q et/ou leurs séquences flanquantes, constituées par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences
20 nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb et/ou une séquence chimère selon la revendication 13,

- l'hybridation dudit support avec au moins une sonde marquée de
25 manière adéquate obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, et

30 - la détection des hybrides formés.

15°) Méthode selon la revendication 14, caractérisée en ce que ledit

transcrit ou ADNc est sélectionné dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

16°) Méthode selon la revendication 14 ou la revendication 15, caractérisée en ce que ledit support comprend en outre toute séquence rétrovirale endogène ou exogène.

17°) Kit de détection et/ou d'évaluation d'une maladie auto-immune et notamment des neuropathologies à étiologie auto-immune, caractérisé en ce qu'il comprend outre les tampons nécessaires à la mise en œuvre d'un procédé selon l'une quelconque des revendications 14 à 16 :

- des réactifs A de diagnostic selon l'une quelconque des revendications 5 à 9, et

- des réactifs B constitués par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb,

lesquels réactifs sont de préférence fixés sur un support approprié.

18°) Kit selon la revendication 17, caractérisé en ce que lesdits réactifs B sont sélectionnés dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

19°) Produits de traduction, caractérisés en ce qu'ils sont codés par une séquence nucléotidique selon l'une quelconque des revendications 1 à 4.

20°) Peptide, caractérisé en ce qu'il est susceptible d'être exprimé à l'aide d'une séquence nucléotidique sélectionnée dans le groupe constitué par les séquences SEQ ID NO:1-22, 28 et 61 selon l'une quelconque des revendications 1 à 4.

21°) Peptide selon la revendication 20, caractérisé en ce qu'il

englobe les peptides dérivés comprenant entre 5 et 540 aminoacides et notamment un fragment de 538 aminoacides, commençant à la première méthionine de la séquence SEQ ID NO:26 (envérine).

- 22°) Peptide selon la revendication 20 ou la revendication 21,
5 caractérisé en ce qu'il est sélectionné dans le groupe constitué par :
- . les séquences SEQ ID NO:23-36 ;
 - . la séquence SEQ ID NO:58 ;
 - . un fragment C-terminal de la séquence SEQ ID NO:26, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première
10 méthionine de la séquence SEQ ID NO:26 ;
 - un peptide de type CKS-17/CKS-25 présent dans l'une des séquences SEQ ID NO:23-36 ou 58 ; et
 - les peptides présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II et notamment les fragments 399-471, 244-271
15 de l'envérine, ainsi que les peptides de séquence SEQ ID NO:68-118, conformément au Tableau I.

- 23°) Peptide selon l'une quelconque des revendications 20 à 22, caractérisé en ce qu'il est obtenu à partir des séquences nucléiques selon l'une quelconque des revendications 1 à 4, dans lesquelles au moins un codon non-sens peut
20 être remplacé par un codon codant pour l'un des aminoacides suivants : Phe (F), Leu (L), Ser (S), Tyr (Y), Cys (C), Trp (W), Gln (Q), Arg (R), Lys (K), Glu (E) ou Gly (G).

- 24°) Compositions immunogènes ou vaccinales, pour la protection contre les maladies auto-immunes, notamment chez les sujets à risque, caractérisée en
25 ce qu'elle comprend au moins un peptide comprenant au moins un motif de type CKS et/ou au moins un motif sélectionné dans le groupe constitué par les peptides présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II et au moins un véhicule pharmaceutiquement acceptable.

- 25°) Composition selon la revendication 24, caractérisée en ce que
30 ledit peptide présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II, est sélectionnée dans le groupe constitué par les peptides tels que

définis dans le Tableau I.

26°) Composition selon la revendication 24 ou la revendication 25, caractérisée en ce que ledit peptide présente la séquence SEQ ID NO:120.

27°) Anticorps, caractérisé en ce qu'il est dirigé contre l'un ou
5 plusieurs des peptides selon l'une quelconque des revendications 20 à 23.

28°) Composition pharmaceutique, caractérisée en ce qu'elle comprend des anticorps neutralisants produits à partir des peptides du Tableau I (SEQ ID NO:68-118) et leurs homologues.

29°) Procédé de dépistage immunologique différentiel de séquences
10 rétrovirales endogènes humaines de la famille HERV-7q normales ou pathologiques, caractérisé en ce qu'il comprend la mise en contact d'un échantillon biologique avec un anticorps selon la revendication 27, la lecture du résultat étant révélée par un moyen approprié, notamment EIA, ELISA, RIA, fluorescence.

30°) Procédé d'identification et de détection de motifs rétroviraux
15 endogènes, anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier autoimmunes, au premier rang desquelles la sclérose en plaques, caractérisé en ce qu'il comprend l'analyse comparée des séquences extraites d'un échantillon biologique avec les séquences selon l'une quelconque des revendications 19 à 23.

20 31°) Application des séquences selon l'une quelconque des revendications 1 à 9, 13, 14 ou 19 à 23 au diagnostic, au pronostic, à l'évaluation de la susceptibilité génétique, à toutes maladies humaines induites, innées ou acquises en particulier celles à composantes cancéreuses, autoimmunes et/ou à incidence neurologique, comme la sclérose en plaques, les syndromes associés et les maladies neuro-
25 dégénératives où intervient tout ou partie des séquences selon l'une quelconque des revendications 1 à 5 et des formes endogènes ou exogènes apparentées.

32°) Séquences nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs selon l'une quelconque des revendications 1 à 9, combinés avec des séquences ou motifs d'origine endogène ou d'origine ou induits
30 de manière exogène.

33°) Vecteur recombinant de clonage ou d'expression, caractérisé en

ce qu'il comprend une séquence nucléique selon l'une quelconque des revendications 1 à 4.

34°) Composition immunogène ou vaccinale, caractérisée en ce qu'elle comprend un vecteur incluant au moins une séquence nucléique codant un peptide tel que défini dans le Tableau I, éventuellement associée à une séquence codant un motif de type CKS-17.

35°) Vecteur de thérapie génique, caractérisé en ce qu'il comprend tout ou partie des séquences nucléiques rétrovirales endogènes de type HERV-7q selon l'une quelconque des revendications 1 à 4.

36°) Vecteur selon la revendication 35, caractérisé en ce que lesdites séquences sont sélectionnées dans le groupe constitué par les séquences SEQ ID NO:2, 20 et 21.

37°) Animaux transgéniques, caractérisés en ce qu'ils comprennent tout ou partie d'une séquence de type HERV-7q (SEQ ID NO:1-22 , 28 et 61).

1/64

CCCTGGGGCGGGCTTCCTTTCTGGGATGAGGGCAAAACGCCTGGAGATACAGCAATTATCTTGCAACTGAG	71	
AGACAGGACTAGCTGGATTTCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACGTCAC	143	
CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCA	215	
AAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCGTGAGAGCACAGCAGGAGGGACAACATCGGGATA	237	
TAAACCCAGGCATTGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTTCCCTTTGTATGGGAGCTGTTTTT	359	région
ATGCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCT	431	répétée
TTTGCTCACCGTCCACCCTGCTGTTTGGCCACCACCGCAGACCTGCCGTGACTCCCATCCCTCTGGATCCT	503	R1
GCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTTGCCGTCCCAATTGGGCTAAAGGCTTGCCA	575	
TTGTTCTGCACGGCTAAGTGCCTGGGTTTGTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTT	647	
TCTTCTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCTTGGAA	719	
CCGTGAGGCCAAGAACTCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAAGCGGCTGTACCATCT	791	
TGGAAGTGGTTTACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCCGGTAACATTTTGGCAACCACGAA	863	
CGGACATCCAAAGTGGTGAATAATTGGACCACTTTCATTGCTATTCTGTCTATCCTTCTTAGAATTG	935	
GAGGAAAATACCGGGCACTTGTGCGCCAGTTAAAAACGATTAGTGTGGCCACCGGACTTAAGACTCAGGTGT	1007	
GAGGCTATCTGGGGAAGGGCTTCTAACAACCCCAACCTTCTGGGTGGGGACTTGGTTTGCTCAAGCC	1079	
AGCTTCCACTTTCAGTTTTCTTGGGGAAGCCGAGGGCCGACTAGAGGCAGAAAGCTGTCTCTGAACCTCC	1151	
GGCAGTAGCCGGTTGAGATCATGGTGTAGCCAGAAGTCTCAACAGTCGCCCATGCATGCACCCCTATCTTC	1223	
CTTCTGACCCATACCTCCTGGGTCCCAACCACAACTTTCTTCAAGTGTAGCCCCAAATTTCTCTTACCTC	1295	
TGAATATACCTTCTCTGATCCCTGCCTCCTAGGTACTATTGGTTCAGACTTCCATTCTCTAGCAAGTTGT	1367	
ATCTCAAAGGGATCTAAGGAAGCTCTGCGCTGCGTCTTAGGCACCTAGGCTATAACCCAGGGAGTCTTAT	1439	
CCCTGGTGTCCCTCCCAATTTAGGCATACAGCTTGTGACATGGGCAGTTATGTAGGACCCACTCCCCACCAC	1511	
CCTTGCCAGGGCCCCAAGTTTGAATAGGCTGAGGGAAGAGAGACAGAGGAGAGAGAGAAATGGAGGA	1583	
GAAAGAGAGAGAGACAGAGAGGAGAGAGAGACAGTGAGAGAGACAGAAGAGAGAGAGACAAAGAGGAGAG	1655	
AGAGAGAGTCAAAGAGAGAAAGAGAGAAAGAAATAGTAAAAACAGTGTGCCCTATTCTTTAAAGCCA	1727	
GGGTAATTTAAACCTGTACTTGAATAATTGAAGTCTTCTCTGTGACCCATAGCAGCTCAATCCACTTTG	1799	
TGGTCAGTGTAAATAAGAGCATAGGCCGAAAGCACTGAGGCCATTGACAACCCGTAGCTTCCCTATCAAAAA	1871	
TCCTTAACCCAGTAACCCGAGATGGACCAATGCATTGAGTCGGTAGGCCAAGTGTCTTGTAAAAGTAGA	1943	
AAAGTAACTTTTAGAGGAAACCTCATTGTGACACACCTCACCTGTTTCTGTAAGAAATTTCTAATAAAAAAGCA	2015	
ACCAGTAACTTTTAGAGGAAACCTCATTGTGACACACCTCACCTGTTTCTGTAAGAAATTTCTAATAAAAAAGCA	2087	
ACCAGTAACTTTTAGAGGAAACCTCATTGTGACACACCTCACCTGTTTCTGTAAGAAATTTCTAATAAAAAAGCA	2159	
ACCAGTAACTTTTAGAGGAAACCTCATTGTGACACACCTCACCTGTTTCTGTAAGAAATTTCTAATAAAAAAGCA	2231	
CACATGGGTATTTCAGTAATTGATACGGGGACTCTTGTGGAAGCAGAGTTAGAAAAATTCCTAATAAACTGG	2303	
TCTCTCAAACGTGTGAGCTGTTTGCACCTCAGCCAAGCCTTAAAGTACTTACAGAATCAAAGACTATCTCA	2375	
ATCTGATTCAAAGGTTAGCTACACCTCTCTGTAATGCATTGCTAAGAACTTGTATGGGAATGCAT	2447	
CTTGATGGGGCAGCTGGGTTGTTATAAAATAGGAACCCAGCCAGCTCTAGGACTCACCCCTGAGCGCAAG	2519	
GCAATGTTGGGCATGCTGGTAAAGGACCACTAGAATCCAGCAGCCAGACCCCTTTCTTTGTGGTCAAGAAA	2591	régions
GGCGGGAAGAGGGGTGACAGGACTGTACATCGGTAAGCATTAATCCGATAAACAGAGGTCCTAGGGTGG	2663	répétées
TTACGACCCCTGGAAGGAACTCACCCCTGAGCACAAAGGCAATGTTGGGCACGCTGAGGACTGCAACATCAG	2735	en tandem
AATCCAGCAGCCTTGACCCCTTTCTTTGTGGTCAAGAGAGGCAGGAAACAGGTGACGAGTGCACATCAG	2807	R2
TGAGCATACTAATTCGATAAGCAGAGGTCATGGGTGGTGTATGCACCTGGAAAGAAATAGCATTAGGACC	2879	
ATAGAGGACACTCCAGGACTAAAGCTCATCGGAAATGACTAGGGTTGCTGGCATCCCTATGTTCTTTTTC	2951	
AGATGGGAAACGTTCCCGAAGACAAAACGCCCTTAAGACGTATTCTGGAGAATTGGGACCAATTTGACC	3023	
CTCAGACACTAAGAAAAGAAACGACTTATATTCTTCTGCACTGCCCTGGCACTCTGAGGGAGTATAAAT	3095	
TATAACACCACTCTTACAGCTAGACCTCTTTGTAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAACT	3167	
TTCTTTTCATTAAAGAGCAACTCAAAATTATGTAAGAAAGTGTGATTATGCCCTACAGGAAGCCCTTACAGAT	3239	
CTACTCCCTTATCCCGACATCCCGACTCTTCCCAACTAATAAGGACCCCTTCAACCCAAATGGTCCA	3311	
AAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAATATTCCTCAATATGACCCCTCCAAGC	3383	
AGTGGGAGGAGAGAAATTCGGCCAGCCAGAGTGCATGTGCTTTTCTCTCCAGACTTAAAGCAAAATAAA	3455	
AACAGACTTAGGTAATTTCTCAGATAACCCCTGATGGCTATATTGATGTTTTACAAGGTTAGGACAATTTCTT	3527	
TGATCTGACATGGAGAGATATAATGTCACTGTAAATCAGACACTAACCCCAATAGAGAAAGTGCCACCAT	3599	
AACCTGACGCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGATAGGATGACAACAGAGGAAAG	3671	domaine
AGAATGATTTCCACAGGCCAGCAGGCAGTTCCTCAGTCTAGACCTCATTTGGGACACAGAATCAGAACATGG	3743	gag
AGATTGGTGTGACAGACATTTGCTAACTTGTGTGCTAGAGGACTAAGGAAACTAGGAAGAAGTCTATGAA	3815	
TTACTCAATGATGTCCACCATAACACAGGGAAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGA	3887	
GGCATTGAGGAAGCGTCCCTCTCTGTACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTAT	3959	
CACCTAGCTCAGCTGCAGACATTAGAAAAAACTTCAAAGTCTGCCGTAGGCCCGGAGCAAAACTTAGAAAC	4031	
CCTATTGAACCTTGGCAACCTCGGTTTTTATAATAGAGATCAGGAGGAGCAGGCCGAACAGGACAAACGGGA	4103	
TTAAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCAAGTGAGACTTTGGAGGCTCTGGAAGGAGGAAA	4175	
GCTGGGCAATTAATGCTTAATAGGGCTTGCTTCAGTGGGCTTACAAGGACACTTAAAAAAGATTGTC	4247	
CAAGTAGAAGTAAGCCGCCCCCTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCA	4319	
GGGGACAAAGGTCCTCTGAGTCAGAAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGC	4391	
AAGCGCCATCCATGCCATCACCTCAGAGGCCCTGGGTATGCTTGACCATTGAGGGCCAGGAGGTTGTCT	4463	
CCTGGACACTGTGCGGTCTTCTAGTCTTACTCTTCTGTCGGGACAACTTCCTCTCAGATCTGTCTACTAT	4535	
CTGAGGGGCTCTAAGACGGGCAGTCACTAGATACTTCTCCAGCCACTAAGTTATGACTGGGGAGCTTTAT	4607	
TCCTTTTACATGCTTTTCTAATTATGCTTGAAGGCCCACTTCTGTTAGGGAGAGACATTCTAGCAAAAG	4679	
CAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTGTGCCCTGCTTGAAGGAAGGAATTA	4751	
ATCTGAAGTCTGGGCAACAGAGGCAATATGGACGAGCAAAAGAAATGCCCGTCTGTTCAAGTTAACTAA	4823	
AGGATTCCACCTCTTTCCCTACCAAAGGCAGTACCCCTCAGACCCAAAGGCCCAACAAAGGACTCCAAAGA	4895	
TTGTTAAGGACCTAAAAGGCCAAGGCCTAGTAAACCATGCAGTAACCCCTGCAGTACTCCAATTTTAGGAG	4967	
TACAGAAACCAACAGACAGTGGAGGTTAGTGCAAGATCTCAGGATTATCAATGAGGCTGTTGTTCTCTAT	5039	domaine
AGCCAGCTGTACCTAGCCCTTATACTCTGCTTTCCCAATACCAGAGGAAGCAGAGTGGTTTACAGTCTG	5111	pol
ACCTTCAGGATGCCTTCTTCTGCATCCCTGTACATCCTGACTCTCAATCTTGTGTTGCTTTGAAGATACTT	5183	

FIGURE 1.1

2/64

CAAACCCACATCTCAACTCACCTGGACTATTTTACCCCAAGGGTTCAGGGATAGTCCCCATCTATTTGGCC	5255
AGGCATTAGCCCAAGACTTGAGCCAACTCTCATACCTGGACACTTGTCTTCGGTAGGTGGATGATTTACTT	5327
TTGGCCGCCATTAGAAACCTTTGTGCCATCAAGCCACCCCAAGCGCTCTTCAATTTCTCGCTACCTGTGGC	5399
TACATGGTTTCCAAACCAAGGCTCAACTCTGCTCAGCAGGTTACTTAGGGCTAAATATATCCAAAGGCA	5471
CCAGGGCCCTCAGTGAGGAACACATCCAGCCTTACTGGCTTATCTCTATCCCAAAACCTTAAAGCAACTAA	5543
GGGGATTCTTGGCGTAATAGGTTTCTGCCGAAATGGATTCCAGGTATGGCGAAATAGCCAGGTCATTAA	5615
ATACACTAATTAAGGAACTCAGAAAGCCAAATACCCATTAGTAAGATGGACAACCTGAAGTAGAAGTGGCTT	5687
TCCAGGCCCTAACCCAAGCCCAAGTGTAAAGTTTGCCCAACAGGGCAAGACTTTTCTCATATGTCACAGAAA	5759
AAACAGGAATAGCTCTAGGAGTCTTACACAGATCCGAGGGATGAGCTTGAACCTGTGGCATACCTGACTA	5831
AGGAATTTGATGTAGTGGCAAAGGGTTGACCTCATTTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT	5903
CTGAAGCAGTTAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCA	5975
CTGCTAAAGGAGACTTGTGGCTGTCAGACAACCTGTTACTTAAATGTCAAGGCTCTATTACTTGAAGGGCCAG	6047
TGCTGCGACTGTGCACTTGTGCAACTCTTAAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAAC	6119
ATAACTGTCAACAAGTAATTTCTCAAACTATGCCACTCGAGGGGACCTTTTAGAGGTTCCCTTTGACTGATC	6191
CCGACCTCAACTTGTATAGTGAAGTTCCCTTGTGAAGAAAGGACTTCGAAAAGTGGGGTATGCAGTGG	6263
TCAGTGATAATGGAATACTTGAAGTAATCCCTCCTCAGGAACTAGTGCTCAGCTAGCAGAACTAATAG	6335
CCCTCACTTGGGCACTAGAAATAGGAGAAGAAAAAGGGCAAAATATATACAGACTCTAAATATGCTTACC	6407
TAGTCTCCATGCCCATGCAAGCAATATGGAAGAAAGGGAATCTTAACCTCTGAGAGAACCACTATCAAAAC	6479
ATCAGGAAGCCATTAGGAATATTATTGGCTGTACAGAAACCTTAAAGAGGTGGCAGTCTTACACTGCCGGG	6551
GTCATCAGAAAGGAAAGGAAAGGGAATAGAAGAGAAGTCCCAAGCAGATATTGAAGCCAAAGAGCTGCAA	6623
GGCAGGACCTCCATTAGAAATGCTTATAAAACAACCCCTAGTATAGGGTAATCCCTCCGGGAACCAAGC	6695
CCCAGTACTCAGCAGGAGAAAGCAATGGGGAACCTCAGGAGACAGTTTCTCCCTCCGGGACGGCTAGCC	6767
ACTGAAGAAGGGAAATACTTTTGCTGCAACTATCCAATGGAATTAATTAACCCCTTCTATCAAACTTTT	6839
CACTTAGGCATCGATTAGCACCCTCAGATGGCCAAATCATTATTTACTGGACAGGCGCTTTTCAAACTATC	6911
AAGCAGATAGTCAGGGCCTGTGAAGTGTGCCAGAGAATAATCCCTGCCTTATCGCCAAGCTCCTTCAGGA	6983
AAGCAAGAAACAGGCCATTACCCCTGGAGAAGACTGGCAACTGATTTTACCCACAAGCCCAACCTCAGGGAT	7055
TTCACTATCTACTAGTCTGGGTAGATACTTTACGGGTGGGCAAGGCGCTTCCCTGTAGGACAGAAAAGG	7127
CCCAAGAGGTAATAAAGGCACTAGTTCATGAATAATTTCCAGATTCCGACTTCCCGAGGCTTACAGAGTG	7199
ACAATAGCCCTGAAGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAACACTCAAGGACATCTAAAAA	7271
ACTGCCCTGAAGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAACACTCAAGGACATCTAAAAA	7343
AGCAAAACCCAGGAAACCCACCTCACATGGCTGCTCTGTTGCCATAGCCTTAAAAAGAACTGCAACTTTC	7415
CCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTG	7487
ACCAAGACAGCCCACTTAGTTGACAGACATCACTCCTTAGCCAAATATCAACAAGTCTTAAAAACATTACA	7559
AGGAACCTATCCCTGAGAAGAGGGAAAGAACTATTCCACCTTGTGACATGGTATTAGTCAAGTCCCTTCC	7631
CTCTAATTTCCCATATCCTAGATACATCTTGGGAAGGACCTTACCCAGTCAATTTATCTACCCCAACTGCGGT	7703
TAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATACCTGCAAGAACCTGA	7775
AAATCCAGGAGACAACGCTAGCTATTCTGTGAACCTCTAGAGGATTGCGCTGCTCTTCAACAACAACC	7847
AGGAGGAAAGTAATAAAATCATAAATCCCCATGGCCCTCCCTTATCATATTTTCTCTTTACTGTTCTTT	7919
ACCCTCTTCACTCTCACTGCACCCCTCCATGCCGCTGTATGACCACTAGCTCCCTTACCAAGAGTTCTT	7991
ATGGAGAATGACGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTT	8063
CACTGCCACAGCCCATATGCCCGCAACTGTATCACTCTGCCACTCTTGCATGCATGCAAACTACTCATT	8135
TTGACAGGAAATGATTAATCCTAGTTGTCTGAGGAGTCTGGAGTCACTGTCTGTGGACTTACTTCAC	8207
CCAACTGAGTGTCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAAGTAATCTC	8279
CCAACCTACCCGGGTACATGGCACCTTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAACCCCT	8351
CCGTACCCATATCTCGCTTGGTAAAGCCTATTTAATACCCACCTCACTGGGCTCCATGAGGTCTCGGCCAAAA	8423
CCCTACTAAGTGTGGATGCTTCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATG	8495
GAACAACCTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTAGGACCTCTGTTTCAATCTGAAATAAC	8567
CCATACCTCAAACTCACTGTGTAATAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTG	8639
GGTAACCTCTCCACAAATAGTCTGCCTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCG	8711
TGTTTGAATGGCTTTCAGAACTTATGTGCTTCTCTCATCTTAGTGCCCTTATGACCATCTACACTGA	8783
ACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTCTTCTTTTGTATAGG	8855
AGCAGGAGTGTAGGTGCACTAGGTATGGCATTTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACT	8927
ATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGCGACTCCCTGGTCACCTTGCAGATCAACTTAACTC	8999
CCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGGGGGAACCTGTTT	9071
ATTTTATAGGGGAAGATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCG	9143
AGATCGAATACAACCTAGAGCAGAGGAGCTTCGAAACACTGAGACCCTGGGGCTCTCTCAGCCAATGGATGCC	9215
CTGGATTCTCCCTTCTTAGGACCTTAGCAGCTATAATATTGCTACTCTCTTGGACCTGTATCTTTAA	9287
CCTCCTTGTAACTTTGTCTCTCCAGAACTGAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAA	9359
GACTAAGATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAAGGCAC	9431
CCCTCCTGAGAAATCTAGCTGCACAACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGTG	9503
TCGGCCAACTCCCAACAGCACTTAGGTTTCTCTGTTGAGATGGGGACTGAGAGACAGGACTAGCTGGAT	9575
TTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTG	9647
CAACTTAGCTCACACCTGACCAATCAGAGAGCTCAATAATGCTAATTAGGCAAGACAGGAGGTAAAGAA	9719
ATAGCCAACTCATCTATTGCTGAGAGCAGCAGGAGGACAATGATCGGGATATAAACCCAAGTCTTCGAG	9791
CCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTCTACTCTAT	9863
TAAATCTTGAACCTGCACTTCTGGTCCATGTTTCTTACGGCTTGAGCTGAGCTTTCGCTGCCATCCACC	9935
ACTGCTGTTTGGCCGACCGCAGACCCGCGCTGACTCCCATCTGGATCATGCAAGGTGTCGCTGTG	10007
CTCCTGATCCAGCGAGGCACCCATTGCCGCTCCCAATCGGGCTAAAGGCTTGCCATTGTTCTCTGATGGCTA	10079
AGTGCCTGGGTTATCTTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTCTCTTCTGTGACCCACAG	10151
CTTCTAATAGAGCTATAACACTCACCAGATGGCCCAAGGTTCTTCCCTTGAATCCATAAGGCCAAGAACCC	10223
CAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCAAGTAACACAACCA	10295
TGAGGGTGCAAAATGCATGGGCCACTAATGGTAGAGCAAGAAACAGAAAGGCCCTGGTCTCGAAGGCATC	10367
AGTGAGCTGAAATGCTGCCCTGGATGTCTATTCTAGGTGTTTTCTGCCTGAAGCAGATTAAACCCCTTT	10439
GTTCACTTCTCAAGTAGGGCTTCTATTACAGCCCAATCAATCCCAACCCAGATGACAT	10500

domaine

env

région
répétée
R1

FIGURE 1.2

3/64

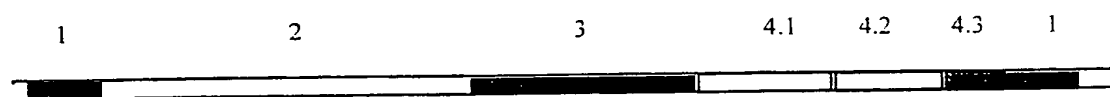


FIGURE 2

[illegible]

FIGURE 3

5/64

IPMALPYHIFLETVLLPSFTLTAPPPCRCMTSSSPYQEFLLWRMQRPGNIDAPSYRSLSKG
 TPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGVTVCWTYFTQTGMSDGG
 GVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEV
 SAQNPTNCWICLPLNFRPYVSI PVPEQWNNFSTEINTTSVLVGPLVSNLEITHTSNLTCV
 KFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPPMT
 IYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALGTGIGGITTSTQFYKLSQELNGDM
ERVADSLVTLQDQLNSLAAVVLQNRRLDLLTAERGCTCLFLGEECCYYVNQSGIVTEKVKKEIRDRIQRRAEELR
NTGPWGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKTYRRPLDRPAS
RSDVNDIKGTPPEEISAAQPLLRPNSAGSS

FIGURE 4

- 1) NSLAAVVLQNRRLDLLTAESGGTFLFLEEK
- 2) NSLAAVVLQNRRLDLLTAERGCTCLFLGEEC
- 3) DSLAAVTLQNHQGLDLLTAEGGGLCYFLGEDC
- 4) DSLAAVTLQNHQGLDLLIAEKGGGLCTFLGEEC
- 5) DSLAAVTLQNCRGLDLLTAEGGGHYTFLGEEC
- 6) LQNRRLDLLFLKEGGLC
- 7) DSLAKVVLQNRRLDLLTAEQGGICLALQEK

FIGURE 5

TSFVEKANGVKCHKYKLSFHKETTHNYVKSVIYALQEAFRVYLPILPASPTPSPTNKDPPSTQMVQKEIDKRVNSEPKS
 ANIPQLXPLQAVGGREFGPVHVFPFSLPDLKQIKTDLGKFSNDPDGYIDVLQGLGQFFDLTWRDIMSLLNQTLTPNER
 SATITAAXEFGDLWYLSQVNDRTTEEREXFPTGQAVPSLDPHWDTESEHGDWCCRHLCTVLEGLRKTRKSMNYSM
 MSTITOGREENPTAFLERLREALRKASLSPDSSEGQLILKRKFITQSAADIRKKLQKSAVGPEQNLETLLNLATSVFY
 NRDOEEQAEQDKRDXXKGHRFSDHPQASGLWRLWKREKLGKLNAXXGLLPVRSTRTLXKRLSKXXAAPSMSPLISRES
 LEGPLPQGTKVLXVRSHXPD/SSSRT

FIGURE 6

TATGAAAAGGCCA
TATGAAAAGAAG
TTTATGTAAAAA
TTTATGTAAAAA
CGACTCCTTCC
:CAACTCCTTCT
:TAACAGTGAAC
:AAACAATGAAC
:CGGCCACGCCA
:-GGCCACGCCA
TTTCTCAGATAA
TTTCTCAGATAA
AGATATATAATG
GAGATATAATG
GAGAGTTTGGCG
GAGAGTTTGGCG
CCACAGGCCA
CCACAGGCCA
GCTGCAGACAT
:GCCACAGACAT
:AATGATGTCCA
:AATGATGTCCC
TTGAGGAAGCG
TTGAGGAAGCA
CTCAGTCAGCT
CTCAGTCAGCT
CCTATTGAAC
:CCTACTGAAC
GGATTAAAAAAA
GATAAAAAAAA
AAAAAGGGAAAA
GAAAAAGGGAAAA
:AAAAAAGATTG
:AAAAAAGATTG
AGGCCCACTGC
AGGCCCACTGC
ACTGAGGCTGCC
ACTGAGGATGCC
GAGGGCCAGGA
GAGGGCCAGGA
CAACTGTCTCC
CAACTGTCTCT

FIGURE 7

7/64

```

01/      TAAATCCCCATGGCCCTCCCTTATCATATTTTTCT
02/      TAAATCCCC-TGGCCCTCCCTTATCATATTTTTCT
03/      TAAATCCCCATGGCCCTCCCTTATCATATTTTTCT
04/      TAGATCCTCATGGCCCTCC-TTGTATATTTTTTTT

01/CTTTACTGTTCTTTTA-CCCTCTTTCACTCTCACTGCACCCCTCCATGCCGCTGTATGACC
02/CTTTACTGTTCTCTTACCCCCCTTTCACTCTCACTGCACCCCTCCATGCCACTGCACCCCC
03/CTTTACTGTTCTCTTA-CCCCCTTTCTCTCTCACTGCACCCCTCCATGCTGCTGTACAACC
04/CTTTACTGTTCTCTTA-CCCCCTTTCACTCTCACTGAACCCCTCCATGCCACTGTACTACC

01/AGT-----AGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGT
02/GTCCATGCCCGTCTCATGCCAGTAGCTCCCCTTAGCAAGAGTTTCTATGGAGAATGCAGCGT
03/AGC-----AGCTCCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTT
04/AGT-----AGCTCCCATACCAAGAGCTTCTATGGACAATGCGGCTT

01/CCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCACTTCACTGC
02/CCCGGAAATATTGATGCCCCATTGTATAGGAGTTTATCTAAGGGAACCCCACTTCACTGC
03/CCCGGAAATATTGATGCCCCATCAAATAGGAGTTTACCTAAGGGAACCCCACTTCACTGC
04/CCTGGAATATTGATGACCCATCGTATAGGAGTTTCTAAGGGAACCCCACTTCACTGC

01/CCACACCCATATGCCCCGCACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTC
02/CCACACCCATATGCCCCGCACTGCTATACTCTGCCACTCTTTGCATGCATGCAAATACTC
03/CCACACCCATATGCCCCGCACTGCTATACTCTGCCACTCTTTGCATGCATGCAAATACTC
04/CCACACCTATATGACCC-----

01/ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCTGGAGGACTTGGAGTCACTGTCTGT
02/ATTATTGGACAGGAAAAACGATTAATCCCAGTTGTCTCTGGAGGACTTGGAG-----
03/ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCTGGAGGACTTGGAGCCACTGTCTGT
04/-----

01/TGGACTTACTTCACCCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGA
02/--GACTCACTTCACTCATACCAGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAACAGA
03/CGGACTTACTTCACCCATACTGGTATGTCTGAGGGGGTGGAGTTCAAGATCAGGCAAGAGA
04/-----

01/AAAACATGTAAAGAAGTAATCTCCCAACTGACCCGGGTACATGGCACCTCTAGCCCCCTACA
02/AAAACATGTAAAGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCCCCTACA
03/AAAACATGTAAAGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCCTACA
04/-----

01/AAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCATACTCGCCTGGTAAGCCTA
02/AAGGACTAGATCTCTCAAACTACATGAAACCTCCATACCCATACTGGCCTGGTAAGCCTA
03/AAGGACTAGATCTCTTAAACTACATGAAACCTCCATACCCATACTTGCCTGGTAAGCCTA
04/-----

01/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAGTGTGGAT
02/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAGTGTGGAT
03/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCTACTAAGTGTGGAT
04/-----

01/ATGCCTCCCCCTGAACTTCAAGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAAC
02/GTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTGTACCTGAACAATGGAACAAC
03/GTGCCTCCCCCTGTATTTCAAGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACAAC
04/-----TGCACCTCAGGCCATACATTTCAATCCCTGTA-----

```

FIGURE 8.1

8/64

01/TCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA
 02/TCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATA
 03/ACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA

01/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA
 02/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACCTCCCA
 03/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCAACTCCCA
 04/-----

01/ATGCATCAGGTGGGTAACTCCTCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTG
 02/ATGCATCAGGTGGGTAACTCCTCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTG
 03/ATGCATCAGGTGGGTAACTCCTCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTG
 04/-----

01/TCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCA
 02/TCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCA
 03/TCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCA
 04/-----

01/TTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTGCATATCTAA
 02/TTCTTAGTGGCCCCCTATGCCCCATCTACACTGAACAAGATTTATACAATCATGTGCATACCTAA
 03/TTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATACAATTATGTTGTACCTAA
 04/-----

01/GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTATAGGAGCAGGAGTGCTAGGTGCAC
 02/GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTATTGGAGCAGGAGTGCTAGGCGGAG
 03/GCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTATCGGAGCAGGAGTGCTAGGTGGAC
 04/-----

01/TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAA
 02/TAGCTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAAACTGTCTCAAGAA
 03/TAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAA
 04/-----

01/CTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC
 02/CTTAAAGGTGACATGGAATGGGTGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTC
 03/CTCAATGGTGACATGGAATGGGTGCGGACTCCCTGGTCACCTTGCAAGATCAACTTAACTT
 04/-----

01/CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGG
 02/CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG
 03/CCTAGCATCAGTAGTCCTTCAAAATGGAAGAGCTTTAGACTTGCTAACCTCTGAAAGAGGGG
 04/-----

01/GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTT-----
 02/GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTT-----
 03/GAAGCTGTTTATTTTTAGGGGAAGAATGTTGTTATTATGTTATTTAGCGGAAGAATGTTGT
 04/-----

01/-----AATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACA
 02/-----AATCAATCCGGAATCATCACCAGAGAAAGTTAAAGAAATTCAGGTGCAATATA
 03/TATTATGTTAATCAATCCTGAATTGTACAGAGAAAGTTGAAGAAATTCGAGATTGAATACA
 04/-----

01/ACGTAGAGCAGAGGAGCTTCGAAA-CACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT
 02/ACGTAGAGCAAAGGAGCTGCAAAA-CACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT
 03/ACGTAGAACAGAGGAGCTTCAAAAACACCAGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT
 04/-----

FIGURE 8.2

9/64

01/GGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA
02/GGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTA
03/GGATTCTCCCCTTCTTAGGATCTCTAGCAGCTCTAATATTGATACTCCTCTTTGGACCCTGTA
04/-----

01/TCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAACTA-----
02/TCTTTAACCTCCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAACTACAAATCGTTC
03/TCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAGTTGTAAAGCTACAAATCGTTC
04/TCTTTAACCTCCTTGTTAAGCTTGTCTCTTGCAGAATCGAAGCTGTAAACTACAAATGCTTG

01/--CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG
02/TTCAAATGGAGCCCCAGATGCAGTCCATGAGTAAATCTACCACGGACCCCTGGACCGGCCTG
03/TTCAAATGGAACCCCAAGATGAAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTA
04/TTAAAATAGAGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTG

01/CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCAC
02/CTAGCCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCAC
03/CTAGCCCATGCTCCAATTGTAATGATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCAC
04/CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCAC

01/AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCC
02/AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC
03/AACCCCTACTATGCCCCAATTCGCGAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCC

04/GACCCCTACTACACCCCAATTTAGCGGGAAGCAATTAGAGCAGCCTATGGCCACCTCCCC

FIGURE 8.3

10/64

CTTCCCCAACTAATAAGGACCCCCCTTTCAACCCAAACAGTCCAAAAGGACATAGACAAAGGA 3
CTTCCCCAACTAATAAGGACCCCCCTTTCAACCCAAACAGTCCAAAAGGACATAGACAAAGGA 4
CTTCCCCAACTAATAAGGACCCCCCTTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAGG 5
CTTCTCCAATAATAAGGACCCCCCTTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAGGG 6
CTTCCCCAAATAATAAGAACCCCCCTTTCAACCCAAACGGTCCAAAAGGAGATAGACAAAGGG 7

GTAACAATGAACCAAAGAGTGCCAATATTCCCTGGTTATGCACCCTCCAAGCGGTGGGAG-- 3
GTAACAATGAACCAAAGAGTGCCAATATTCCCTGGTTATGCACCCTCCAAGCGGTGGGAG-- 4
GTAACAGTGAACCAAAGAGTGCCAATATTCCCAATTATGACCCTCCAAGCAGTGGGAGGA 5
GTAACAATGAACCAAAGAGTGCCAATATTACACGATTATACTCGCTCCAAGCAGTGGGAG-- 6
GTAACAATAACCAAAGAATGCCAATATTCCCCGATTATGCCCCCTCCAAGCGGTGGGAG-- 7

A-AGAATTCGGCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCAC-ACTTGAAGCAAATTAAA 3
A-AGAATTCGGCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCAC-ACTTGAAGCAAATTAAA 4
AGAGAATTCGGCCAGCCAGAGTGCATGTGCCTTTTTCTCTCTCCAG-ACTTAAAGCAAATAAAA 5
-GAGAATTTGGCCAGCCAGCGTGCATGTACCTTTTTCTCTCTCAG-ATTTAAAGCAAATTAAA 6
-GAGAATTCGGCCAGCCAGAGTGCACGTACCTTTTTCTCTCTCTAGACTTTAAA----TTAAA 7

ATAGACNTAGGTNAATTNTCAGATAGCCCTGATGGYTATATTGATGTTTTACAAGGATTAGGA 3
ATAGACXTAGGTXAATTXTCAGATAGCCCTGATGGXTATATTGATGTTTTACAAGGATTAGGA 4
ACAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGA 5
ATAGACCTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGA 6
ATAGACCTAGGTAAATTCTCAGATAACCCTAATGGCTATATTGATGTTTTACAAGGTTTAGGA 7

TTCCTGAGTTCTTGCACTAACCTCAAAT 1
CAATCCTTTGATCTGACATGGAGAGATATAATATTACTGCTAAATCAGACGCTAACCTCAAAT 3
CAATCCTTTGATCTGACATGGAGAGATATAATATTACTGCTAAATCAGACGCTAACCTCAAAT 4
CAATCCTTTGATCTGACATGGAGAGATATAATGTCACTGCTAAATCAGACACTAACCCCAAAT 5
CAATCCTTTGATCTGACATGGAGAGATATAATGTTACTGCTAAATCAGACACTAACCCCAAAT 6
CAATCCTTTGATCTGATATGGAGAGATATAATGTTACTGCTAAATCAGACACTAACCCCAAAT 7

GAGAGAAGTGCCGCCATAACTGCAACCCAAGAGTTTGGCGATCCCTGGTATCTCAGTCAGGTC 1
GAGAGAAGTGCTGCCATAACTGGAGCCCGAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC 3
GAGAGAAGTGCTGCCATAACTGGAGCCCGAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC 4
GAGAGAAGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTC 5
GAAAAAAGTGCTGCCATAACAGCAGCCTGAGAGTTTGGCGAACTCTGGTATCTCAGTCAGGTC 6
GACAGAAGTGTCGCCGTAACCTGGAGCCCGAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC 7

AATGACAGGATGACAACAGAGGAAAGATAATGATTCCCCACAGGCCAGCAGGCAGTTCCAGT 1
AATGATAGGATGACAACGGAGGAAAGAGAACGATTCCCCACAGGGCAGCAGGCAGTTCCAGT 3
AATGATAGGATGACAACGGAGGAAAGAGAACGATTCCCCACAGGGCAGCAGGCAGTTCCAGT 4
AATGATAGGATGACAACAGAGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCAGT 5
AATGATAGGATGACAACAGATGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCAGT 6
AATGATAGGATGACAACAGAGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCAGT 7

GTAGACCCTCATTAGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTGCTAACT 1
AAT 2
GTAGCTCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTACTAACT 3
GTAGCTCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTT 4
CTAGACCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCTGCAGACATTTGCTAACT 5
GTAGACCCTCATTAGGACACAGAATCAGAACTTGGAGATTGGTGCCACAGACATTTGCTAACT 6
GTAGACCCTCACTGGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTGCTAACT 7

FIGURE 9.1

11/64

TGCGTGCTAGAAAGGACTAAGGAAAAC TAGGAAGA----	1
TGCGTGCTAGAAAGGACTAAGGAAAAC TAGGAAGA---CTATGAATTATTCAATGATGTCCACT	2
TGCGTGCTAGAAAGGACTAAGGAAAAC TAGGAAGA---CTATGAATTATTCAATGATGTCCACT	3
TGTGTGCTAGAAAGGACTAAGGAAAAC TAGGAAGAAGTCTATGAATTACTCAATGATGTCCACA	5
TGCGTGCTAGAAAGGACTAAGGAAAAC TAGGAAGAAGCCCATGAATTATTCAATGATGTCCCCT	6
TGCGTGCTAGAAAGGACTAAGGAAAAC TAGAAAGAAGCCTGTGAGTTATTCAATGATGTCCACT	7
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	1
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	2
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	3
ATAACACAGGG-AAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	5
ATAACACAGGG-AAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAAGGATTGAG	6
ATAACACAGGG-AAAGGAAGAAAATCCTACCGCCTTTCTGGAGTGACTAACGGAGGCATTGAG	7
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAAAGGGAAAAGTTGGGAAAAGTA	1
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAAAGGGAAAAGTTGGGCAAATTG	2
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAAAGGGAAAAGTTGGGCAAATTG	3
GAAGCGTGCC232AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG	5
GAAGCATACC238AGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGAAAAGCTGAGCAAATTG	6
GAAGCATACC233AGGCAAGCGGACTTTGGAGGCTGGAAAAGGGAAAAGCTAGGCAAATCA	7
TATGTCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	1
AATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACGCTTTAGAAAAGATTGTCC-AA	2
AATGCCTAA	3
AATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	5
AATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	6
AATGCCTAATAGGGTTTGCTTCCAGTGCAGTCTACAAGGACACTTTAAAAAAGATTGTCCAA	7
-TAGAAATAAGCCACCACCTCGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	1
GTAGAAATAAGCCGCCCC-TCGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	2
GTAGAAGTAAGCCGCCCCCTCGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT	5
GTAGAAACAAGCTGCCCCCTTGTCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	6
-TAGAAATAAGCCGCCCCCTCGTCCATGCACCTCGTGTCAAGGGAATCACTGTAAGGCCCACT	7
GCCCCAGGGGATGAAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	1
GCCCCAGGGGACGAAGGTCTCTGAGTCAGAAGCCACTAACCTGATGA	2
GCCCCAGGGGACAAAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	5
GCCCCAGGAGATGAAGGTCTCTGAGTCAGAAGCCACTAACCAGATAA	6
GCCCCAGGGGACGTAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	7

FIGURE 9.2

[illegible]

FIGURE 11

[illegible]

FIGURE 12

15/64

agttgcaattccttgctcaactctgagagaaaccccagccacatctccagcaaacaaga
|||||
agttgcaattccttgctccactgtgagacaaaccccagacacatctccagcacacaaga 2299

acttcaaaacacctgaactgcagcagccaggcggttctccaggaccacctccccaggat
|||||
acttcgaaatgcctcaacctcaggtgccagggggttctccagaaccttctccccaggag 2359

cttgcttcaagtgccggaaatctgaccattgggccaaggaatgcctgcagcccaggattc
|||||
cttgctacaagtgccagaaatctggccactgggccaaggaatgccacagaccaggattc 2419

ctcctaagccacgtcccatttgtgcaggacccactggaaatcggactgtccaactcacc
|||||
ctcctaagctgtatcccatctctgtgggacccactaaaaatcagactgttcaactcacc 2479

cggcagccaatcccagagcccctggaactctggccaaggctctctgactgactccttcc
|||||
tggcagccacttccagagcccctggaactctagcccaaggctctctgactgaccttct 2539

cagatcttctcggttagcagctgaagactgacactgcccgatcacttcagaagtccct
|||||
gagatcttcttggttagcagctgaagactgacactgccagatcgctcggaagcctaca 2599

ggaccatcacggatactgagcttcaggtaactctcacagtggaggctaagtccatccct
|||||
ggaccatcacagat-----gctccaggtaactctcacagtagagggttaagtctgtccct 2654

gtttaatcgatacaggggctacccactccacatcaccttcttttcaagggcctgtttcc
|||||
tcttaatcaatatggaggctacccactgcacattaccttcttttcaagggcctgtttcct 2714

tttccccataactgttggtgggtattgacggccaagcttcaaaaccccttaaaactcccc
|||||
ttgctccataactgttggtgggtattgacggccaggcttctaaacctcttaaaactcccc 2774

cactctggtgccaaacttggacaacattcttttatgcactctttttcagttatcctcacct
|||||
aactctagtaaccaacttagacaatactcttttaagcactccttttttagttatccacctt 2834

gcccagttcccttattaggccgagacattttaaccaaattatctgcttccccgactattc
|||||
gcccagttcccttatgaggccgagacacttcaactaaattatctgcttccctgactattc 2894

ctgggctacagccacatctccttgccgccttcttcccaacccaagcctccttcatac
|||||
ctggactacagctacatctcattgctgccccttcttcccaatccaagcctcctttgcatc 2954

ttcctctcatatccccccaccttaaccacaaagtatgggacacctctactccctccctgg
|||||
ttcttgt---atcccccaaccttaaccacaaagtataagataacctctattccctccttgg 3011

FIGURE 13.1

3071

3131

3191

3247

3307

3367

3427

3487

3547

3607

3660

3720

3778

3838

FIGURE 13.2

atataaactcacaaaaggaaacctagctgaccccatagattcctaaatcctttcccccactc
||||| 3898
atataaactcacaaaaggaaacctagctgaccccatagattcctaaatcctttcccccactc 3898
ctctttccattccttgaagacagcttttagagactgctccacactagctctccctgtctc
||||| 3958
ctctttccattccttgaagacagcttttagagactgctccactctagctctccctgactc 3958
atcccaaccctttttcattacacacagccgaagtgcagggctgtgcagtcggaattcttac
||||| 4018
atcccaacactttttcattacacacagctgaagtgcagggctgtgcagtcagaattcttac 4018
acaaggaccgggaccatgcctgtagcctttttgtccaaacaacttgaccttactgtttt
||||| 4078
acaaggaccgggatcgcatcctgtagcctttttgtccaaacaacttgaccttactgtttt 4078
aggctcgccatcatgtctccatgcggtagcttccgctgccctaataacttttagaggcct
||||| 4138
aggctggccatcatgtctccatgcagcgtctgctgccaccctaataacttttagaggcct 4138
caaaatcacaaactatgctcaactcactctctacagctctcacaacttccaaaatctatt
||||| 4198
caaaatcacaaactatgctcaactcattctctacagctctcataatttccaaaatctatt 4198
ttctttctcacacctgacgcataactttctgtctcccggtccttcagctgtattcact
||||| 4258
ttcttctctcacacctgacacataactttctgtctcccggtccttcagatatactcact 4258
ctttgttgagtctcccacaattaccattcttctggtgccagacttcaatctggcctcca
| 4316
c--catttattctcccacaattaccattattctggtgccagacttcaatccggcctcca 4316
cattattctggataccacacctgacctgatgattgtatgtctctgatctacctgacatt
||||| 4376
cattattctggataccataacctgacctcatgactgcatctctctgatccacctgacgtt 4376
caccctatttccccatatttcttctttctgttctcatgttgatcacatttggtttac
||||| 4436
caccctatttccccacatttcttctgacctgtttctcaccctgatcacacttggtttat 4436
tgacggcagttccaccaggcctgatcgccactcaccagcaaaggcaggctatgctat
||||| 4493
tgatggcagttccaccaggcctaatacgccactcaccagcaaaggcaggatagctat 4493
gaactgattgccttaactcgggccttcactcttgcaaagggaactacacgtcaatatattat
||||| 4553
gaactagttgccttaattcaagcctcactcttgcaaagggaactacgtgtcaatatctat 4553
actgactctaaatatgccttccatatcttgaccaccatgctgttatatgggctgaaaga
||||| 4613
actgattctaaatatgccttccatatcttgaccaccatgctgttatatgggctgaaaga 4613
ggtttccctcactacgcaagggtcctccatcattaatgcctctttaataaaaaactcttctc
||||| 4672
ggtttccctcactacacaagtgtcctccatcattaatgcctctttaaaaaa-ctctqctc 4672

FIGURE 13.3

18/64

aaggctgctttacttccaaaggaagctggagtcacacactgcaagggccacaaaaggcg
|||||
aaggctgctttacttccaaaggaagctggggtcattcactgcaagggcatcaaaagact 4732

tcagatcccattactctaggaaatgcttatgctgataaggtagctaaagaagcacctagc
|||||
tcagatcccattgctctaggcaatgcttatgctgataaggtggctagacaagcagctagc 4792

gttccaaacttctgtccctcatggccagtttttctccttcccatcagtcattccacctac
|||||
tctccaaacttttgtccctcatggccagtttttctccttcacatccgtcactccacctac 4852

tccccattgaaacttccgctatcaatctcttctcacacaaggcaaatggttcttagac
|||
tccacagctgaaacttccacctatcaagctcttcccccgcaaggtaaatggttcttagac 4912

caaggaaaatatctccttcagcctcacaggccattctattctgtcatcatttcataac
|||||
caaggaaaatatctccttcagcctcacaggccattctattctgtcgtcatttcataac 4972

ctcttccatgtaggttacaagccactagtcacactcttagaacctctcatttcctt-cca
||
cttttccatgtaggttacaagccactagcctgtctcttaggacctctcatttcctttcca 5032

tcgtggaaacatatcctcaaggaaatcacttctcagtgttccatctgctattctactacc
||
tcatggaaatctatcctcaaggagatcacttctcagtgttccatctgctattctgctacc 5092

cctcagggattgttcaggccccctccctccctacacatcaagctcggggatttgccct
|||||
cctcagggattgttcaggcctcctcccttctacacataaagctcggggatttgccct 5152

gccaggactggcaaattgactttactcacatgccttgagtcaggaaactaaaatacctc
|||||
gccaggactggcaaattgactttactcacatgcctcgggtcagaaaactaaaatatctc 5212

ttggtctgggtagacactgtcactggatgggtagaggcctttccacaggggtctgagaag
||
ttagtctgggtagacactttcactgggtgggtagaggcctttcccatagagtctgagaag 5272

gccactgcagtcatttcttcccttctgtcagacataattccttgggttggccttcccacc
||||
gccaccgcggtcatttcttcccttctgtcagacataattccttgggttggccttcccttc 5332

tctatacagtccaataacggagcagcctttattagtcaaatcacctgagcagtttttcag
|||||
tctatacagtctgataacggaccagcctttactagttaaatcacccaagcagttttctcag 5392

gctcttgggtattcagtggaaaccttcgtacccttactgtcctcaatcttcaggaaaggta
|||||
gctcttgggtattcagtggaaaccttcatacccttaacatcctcaatcttcaggaaaggta 5452

gaatggactaatgggtcttttaaaaacacacccccaccaaactcagcctccaacttaaaaag
||
aaaccgactaatgggtcttttaagacacacctcaccaagctcagcctccaacttaaaaag 5512

FIGURE 13.4

19/64

TGCCTTTATTTCCGTAGGCTGGTCATATGGCGCTAGCACTCACATAAAGCTACCGAGGAG
 AGCGAATGAAACCAAATCACTTTACCTTCACAGCACGAGGCCGTCGTCCCTCTCGATAT
 TTGGCCCCGTGTGTGCGATACCGCCCTCTGGACGTGGTGATCAAATAAACTCCCTAGCTCC
 CCGCCGCTCGACGCCATCTTGCTACTTTGATCCTCGCAGGGAGGACAACATCCGCCCTA
 CTGAGCTCCCTTTTATCCAATAAGAGAGCGGGATGAGTTAAGGAGTGCCAGGATTGGCTG
 GAGAATCGACAGCGTCGGCCATCGTTTTCTGCGTGCGAAGATTTGATGAACGAGGTGCCG
 CCCCCGAGCGGCTCGGCGGAGAGGCGCGGTGGGTGACAGAAGCTTTCTTGTCACCCAC
 TACAGGCTTACGGCAGGATGCGCAGCGGGGAGAGGGGGCGGGGCCGAGGGGGCGGGGCC
 GATCGATCTCCTCCGGCTCCGACGTCTCGGCCTGCGGGGTCCCGGGTCCCTTTGCGGCGC
TAGGGTGGGCGAACCAGAGCGACGCTCCGGGACGATGTGGGGCAGCGATCGCCTGGCGG
GTGCTGGGGGAGGCGGGGCGGCAGTGACTGTGGCCTTCACCAACGCTCGCGACTGCTTCC
TCCACCTGCCGCGGCGTCTCGTGGCCAGCTGCATCTGCTGCAGGTAACCTGCCGGCCCC
 GAGCCACCTGATCTTCAGCCTGGGTGCGACGAGGCCGAAGCCTCTCAGGGACGCGGCGG
 GACACCGGCTGCCACCCGGGCGCCGCCGAAGCGCGCAGAGATCAGGGTCCCTCGACGGCA
 GGGCCCTTCTGGGTAGTCTCTGGATCCCAAGTCCAGTGCAGCCCTGGGCTCGTCTTAT
 CCCAGGTCTTTTCACTTGGTGAACTGAACCTAGAAACGTCCTAATATTCTACCACTGTT
 TTTATAAATATTCCTTATTCCAGGCTGAAAAGCTCCTGAGAAGTGGTTTGTTTTTATTA
 TTTTAAAAGGTGTTTTCTTGCCAGCCATTTCCAGTTAACCTGCGCTGCTGCCGTCCGGG
 CCGCGAGAGCGGGACGCAGAGTTGTTGGCGGAGCCCTGTCGGTTCCCGGGGACTAAGCA
 CCGCGTCCCATGAGCGGGAAAGGTTAATACAATGATGGTTCTGCCCTGCGTCGCTGACGC
 GGAACACAGCTGTAGTGTGTTAGGAACACATAACGTAGTTAAGATCACTTGAAGCTCTGC
 GATCAGTCGCCCCTTCTGGACGTTGTGGTTAGGATGTTTCACAGTTCTAACCCTGGTGGA
 GATACAGCGTCCATATTTTCATAATTAATAAGAGGCACATGGTCTCACGAGTTTGAGT
 GTACTTATGGGGGCAAAGGACGGCGTATTTGAAATCCTCATAAATCCTGGATGCATGGT
 ACCCACCAGTGGCTAATCTATGCAATGAATAGAGTTTGCAATAATTTCAAGCATCCCTTC
 TTTCCACTTGAGTTACTTCCCCATACCTAGGGGAAGATATTTTTTGGTCCACTGAAAACAT
 GAGTTCAGCAGAATCCTCCTATCATCGTCGTTATTATTTTTTACCCTAAGTAGACAATC
 TTTTGGTTTTTGATGGGCTTTATGGCTAGAGACAAATCAGTCACGTGCACCAAGTCCAG
 GTAGAAGTTGGTTCAGTGCTCTGTGACGCTTCGATGGGATTTTTTCAACATGTTTTCAAATC
 TGCACCTTAATAGTAGGAATGCTTTCTTACAGTAACCTCTAATTTGATCCTAAGATGTAGTT
 GTTACCTTACATTCATCACTGTTTAAAGAATTTAGTGGTCTTGATCTTTGTTTTAAATTTT
 GAGCCTTCGGGAAGTACTTATAAGAAATTAATTCATGCATATCTTTTTGAAATGTAAATGT
 CTTTAGCCCTGGAACAAATTGCTGTTTCTGTTGAGCCCATATTAGCAGAATAGGTCAACT
 TTACTTTCTAATTTATCAATGTAATAAGTTTATTACTTTATAGATTCCATAAATCTATACA
 TTTATTCCTCGATGAATTTATATAAATTTATAGAATTTATGTTTTATAGAAAATTTGGA
 GCATGGAAAATTTATTAACAAGAAAATAAGTTACCCATAATCCCAGAACCTTAGAGGTGACT
 AATGTTGACAGTTTGGATCAAATCTTCAGTTTTGTTTCTAATCTTTATTTTTAACATAA
 ATGAGGTCTGTATACACACGTACAGTTTTGTGTCCTGGTGTTTTTATTTAATGTTATTA
 TGAGTGTTTTATTTTTGTTAAAAGGTCATCATTTTAAAGTTGTTAATTAGTATTCTAGCACA
 AATTTGCCATAATTTATTTAATTTGTTTACTATGATTGACCATTTAGATTGTACTTAATTT
 TTAGGCATTAGAAGTGATAAACTATATTTTAAATCAGACGTTGAAAATAACACATCTTTGT
 TTAGAAAACATCATTTTATTTCTGGTTGTCTAGGATAGATTCCCAGAATTCTTGGGTTAG
 AGGCCATAGATAAATTATGAAAGCAGAAAGATTACAAGTTGGGAGTTAATACTTGAATTA
 CTTTATTTGGGGTGAAGCATTGAGTGCATAATACAGATCATGCAGTAATGGGAAGAAGGG
 TTGGAACAATGGTTTTCTGGCCTATGTGACACTTACCTTGAAGCTTTTAAAGATACAGAT
 GTTCTGATCAACCCTCAGACCTATTAATCAGACCTAAAATCTTAGGGGAATAGGCTTTAG
 GCATCTCTAATTTTAAAAAATTTATTCAGGCTACTTGGATGCACAAAAGAGTTGAGACCT
 ACTGTCCTAGAAATCATAGAATTTTAAATGACGATAGAGACCTTAAGCATCTAGGTCGTTTC
 TGACTTTTACATGTAAGGAACTGGCATTCTTAGGCCAGTACCATTGCCATGCAGCTAA
 TTTGCCCTCTTGCTATAGCTCACTCTGCATCACCAACCTACCGTTCTCACTGTTTCTT
 CTATAACCAATCTCCTTCCCACTTCTGTTCTCTTACTCATGCCATTCTTCCCTCAGTCAT
 TTTTCTTCTTCCATACAAATTCATGTCTTTAAAAGGAATAATCCTACCTCCTCCACA

FIGURE 14.1

20/64

TAGCTTTCCAATTCTCTGTGGCCACATTTGTCTCCCTTTCAATACTTCTCTGTTGTGTT
 ATGTGACACATCACATTTGATATACTCTGTACTGTGTTTCAAGTATTGTATTCTCTGTT
 TACTCAAGTCATTATTTTCAGGACTGACTACCCAGTAGATGCTTTAAGTCAGGATTTCTCA
 ACCTTGGCACTGTTGACATTTTGAGCTGGATAAATTTTTGTTTTGGGGGCTCTCCTGTAC
 ATTTTAAAGATGTTTAAACAGCACCTTGGCCTCTATCCAGTAGACGCTGTACTGCCTCCC
 CCTATCTGTGACAACCAAAAGGTCTTCAGACATTGTGAGATGTCTACTGAAGGACAAAA
 TCACCTCTGGTTGAGAACCACCGCTTCAACTAAGTTATCTTCTCTGTACTCAGAACTTGA
 TGTGATTGCAGCAGGGGGAGAGGATTATATACACAGTGAATGCAAACGAACCTAAATCA
 CCATTCCGATATGGCCACACAATTTTCATTTCCCTTGTGTTAGCAAGAGATACCCTAGGC
 TTTGGACCTGATTATTCCTAAGGCATTCTGATGTATGGTTTTACCTGCAGATTTCTGGT
 AATACTGATACCTCAGTTTGGGTCAAAGAAGGTCAATTAATTGATTGATTGATTGACT
 CCTGGAAGACGCTCCTTTCTAGCTGTCTCTTTCTCTCTTACCTGAATAGCCAGGGC
 TCTGTGGTTCAGTGAAGTATTTTGACATAAAAAATTAACCTAGAACATTGGTCTGCAGAG
 TTTGCTCAATATAACTGAGCACATATTGTGGCTTTATGGAGCTGGTTACTACTTTTTGAC
 CAAATAAAATAATTAGAAGTATTTTCTCCTCAATAAGGTTTCAATTTTCTTTTTTTCAGT
 GAGCTGGTAGAGTTTCTTTTTTGATATTTTCAGGGCATCTTTCATATTTCCATCTCTTAA
 GTTCTTTCATATGAAGTAGAATTTATCTGGATTATGTATTGCTGACTCTGATGAAAACCC
 ATAGAAAGCATCTGGGGCTTGATCACCTTCATTCTTGTAAATAGCTCACACGGTTACAGCT
 GATATGGTAACTTAAGACTTTTGATTCAAATCTAGGCAAAATACACTCAGTTGAAAGAA
 TTTGTCAGCCAGAACAGTTGGACTGTTCTGTGAAAATTGTGAGAAAAATTACACAATAA
 GTGATACATGATGATGGCTTTCTTAAATATAAAATTGTAATAACATGGTTAATTTCCAGT
 ACGTTATATTGTCCCAGAAGTGGCTCCAACATTGTTTGAAATTTGTCTCATTTAAAGAAA
 CATAAGCTGGCTATGGTGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCTGAGGCAGG
 CAGATCACCTGAGGTGAGGAGTTCGAGACCAGCCTGGCCAACATGGTAAAACCCCATCTC
 TACTAAAAATACAAAAATTAGCCGGGCATTTGGTGGGGGCTGTAATCCAGCTACTTGG
 GAGGCTGAGGCAGGAGAATTGCTTGAATCTGGGAGGTGGAGGTTGCAGTGAGCCGAGATT
 GTGCCACTGCCCTCCAGCCTGGGTGACAGAGTGAGTCTCCGTCTCAAGAAAAAAAAAAAA
 AAAAGCAAGAAACATAAAGACTGGGCATGTTGGCTCATGCCTGTAATCCAGCACTTTGA
 GAGACTGAGGTGGGAAGATCACTTGAGCCAGGAGGTTAAGGCTGCAGTGAGCCGTGATT
 TTGCCACTGTACTCGAGCCTGGGCAACACAGTGAGATCCTGTCTCAGGAAAAAAAAAATT
 GCATGTAAATGAATGAATTTGATATTTAATATTTTAAATTATGAAAACCTGTTCTGTAGAG
 ATGTAGATCTTGCCATGTTGCCAGGCTGGCTTTGAACCTCTGGGCTCAAACAATCCTCC
 TGTCTCAGTCTCCCAAAGTATAAAGATTACACATGTGAGCCACTGCACCTGGCCTAATAT
 TTTTAACTTAATGAATTTATTTTGATATAAATAAATTAATAACACTGAAGCTTCTTGATA
 TAATAAGTCTTTTTGTGTGTGTGACGGGTTCTCACTCTGTTGCCAGACTGGAGTGTAAT
 GGCATATCATGGCTCACTGTAGCCTCAACCTCCCTGACTCAAGTGATCCTCCACCTCG
 GCTTCTGAGTAGATGGGACCACAGGCGTATGCCACCACACCTGGCTGATTTTTAAAT
 TATTATTGATACATATTAATAAAATTATTTTATTTTAAAAATGATATATGTGGCTGGC
 ATGGTGGCTCATGCCTGTAATCCCGACAGTTTGGGAGGCCGAGGTGGGAGGATCACTTGA
 GACCAGGAGCTTAAGACCAGCCTAAGCAACATAGTGAGATCCCATCTCTATAGAAAAAA
 AAATGGCTAGGTGTGGTGGTGTATGCCTATATTCCAGCTACTCAGGAGACTGAGGTGAG
 AGGATTGCTAGAGCCCAGGAGTTTCAAGTTACAGTGACCTATGATTGTGCCAGTGCACTC
 CAGCCTGGGCAACAGAGCAAAATCCTGTCTCAAAAAAAAAAAAAAGTTGCAAAATGCTTAT
 GATGCAATATAAGTAGTGGAAAAGGATATTAAATTGTGCCTATATGAACACAATAATATG
 AAAAATCTGCACATAGAGAAAAGGATTAACAAGAAATAGACCAAATTGTTACATGGTTG
 TCTTGTGTGTGGAGAGAATATCAGTAGTTCATTTGTTTCTTCCAAGTTTATATGTTTC
 CGAGGTCTCTATAATGAGTTTGTAAATTGTTTAAATCATAGAAAACCTTTTTTGGTCTTG
 GCCACAACTTACATGTTTTAATGTAATTGCTTTTTTAAATGAGAATAAATGTTATATTTT
 GCTTTTTTAAACCTATATTCCCATAGTTATATGAGCCCTTACAATTATTAAGAGGCTGC
 ATAATATAACGTTTCTGGAAGGGTACAGAAGAAACAGCAGTAATTACCTCTGAGAACAGA
 GACATGGCTTCACATTTTACCCTTTTGTACGTTTTGTGCTTTTGCCACATGCATTTATTA
 TTCTTCCAATAAATAAGTAAATAAATATGGATTGTATACTCCATCTGGTTGGTGTTCAT
 AATTCTAAATTTATATTGCTACATTTTTAAAGATGATATGTGTTTCTACTTATTAACGTA

FIGURE 14.2

21/64

TATGTTAAAAATAGTAAATTTATATCTTATTTAATAATTTCCCTATTGATAGACATTTAAG
 ACAGTCTCAAGTGTTCACATCATAGAAAATACTGCACAGATAGCTTTTGCTATAGTTTC
 TTTTTCTTTGAATCGTTAATTGGGAATAAATGCTCAAATAGTTATATGTGGCTCAACTG
 CTATTTAAGTTTATTGACTGACTGCTGCCATTTTGAATTCCTGAAGGGGTTGATTAAATTT
 ATAATGCTGCCATAAGAATATAAGGGTATTGGCTTCATTAGCATCCACCAGCATTGGGTG
 TTGGAAATGATTATAGATTTTTAAATGCTACAACAAATGTAGATAACAGAGAACTATCTA
 TAGAACTCTTTTTGGACATGTGAATTGTAATAATAGTTTATTTTCATGTGAATCCAGAAA
 AATGTATACGAAAACCTTTTTCTCTCATTTCTTATATGAATAGAATCAAGCTATAGAA
GTGGTCTGGAGTCACCAGCCTGCATTCTTGAGCTGGGTGGAAGGCAGGCATTTTAGTGAT
GGGGGACAGGTAAGCACATGTGATGGCAATAACTTTCTTCTAATATCACATAATATAGCA
 ATAGAAATAAAATTTAAAGTTTAGATTTTTTGTAAAGGAGGTGAGATGTCACCTAATTT
 GTATGCTATTATGTAAC TAGTCTAGGATATTGAAGCTGACTATACTCTGTTTTTAGGTCA
 TTATCTTGTAGTTTACCATACTCCCTACTTGCTTCTTATTCTACTATTTAACTCATTTTC
 CACATCCCCTAATTTTGGTTTCATGAAATTATTTTTCTTCTGAATTACTAGGTTCTACT
 TACTATTATTAACTTTTATTCTGACATATTTTATAACCTTCCATGGTCTCACTTGATTA
 AAAATAAAAAAATTCAGCTGGGTGCGGTGGCTCACACCTATAATCCCAGCACTTTGGGAGG
 CCAAGGTGGGCGGATAATTTGAGGTGAGGAGTTGGAGACCAGCCTGCCCAACGTGGTGAA
 ACCCCCCCTCTCTACTAAAAATTCAAAAATTAGCTGGGCATGGTGGCAGGTGCCTGTAAT
 CCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGTGGAGGTTGCA
 GTGAGCTGAGATTGCACTGCTGCACTTCAGCTGGGTGACAAGAGCGAAAACAATGTCTTGA
 AAAAAAATAAAAAATAAAAAATTCTACAACACAGGGTTATTATTTTTCCATTTTTGTTTT
 CCCTTATGAGTTTAATATGTTTAGATTATAAACCTGAAAGCTTGAATACCTATGTCTATC
 TTTTGTTTTCTTATGTTTATCAAGTTATTCCTTTAAACATTTTCTAACTGTAAGAATAA
 TGTGAGGCTGGGCTCAATGGCTTATGCCTGTAATCCAGTGCTTTGGGAGGCCAAGGTGG
 GAGGACCATTGAGGCCACGAGTTCAAGATTAGCCTGGCTAGGCAACATAGCAAGACCTT
 ATCTCTATAAAAAAATTAAAAAATTAGCTGGGCATGGTAGCAAATGCTTGTAGTCCCAG
 CTACTCAGCAGACTGAGGTAGGAGGAATGCTTGAAGCCAGGAATTTGAGTGACCTATGAT
 TATGCACTCCAGCCCCGGGCAATAGCAAGACCTATCTCTTAAAAAGAAGAAGATGTAGTAA
 TAATACATATTTCATTATAACTATTTTACCATTGAAAGTAAAAAATGAGTTTTTACCTTTT
 CCCAGTCCCATCCTCAGAATGGGGATCTCAGTAGACCTTTAGGATTGGAAGAATGAGATC
 ATTCATATTTTCTGCAATTATTACCCACAAAAATATTCAGATACCTTTCCATGTATTAC
 AAACAATGTGCATTTAACATGTCTCTCTCTCTCTCTCTCTGTGTGCGTCTTCATGA
 TCCTCTGTTGAGCCCTGCCAGTAAGACACTATCTCCTGAAGAATCACTGATAGGAACAG
 AAAGTGGACTGGCTAGGCCAGGAGTCCTTAGCTTCTTAGGGGGCAGGAGCTGCTTTGTGC
 TTTCTCAGAATCAGATATATATGTGGACTGAAACATTTAAAAACAGAATAGCCAAGGGTG
 CTATACGTTTTAAACTTATATAGATGGGGCTACATTGCTCTCTATTACTAATTTCCCATG
 ACAATACACGAGAGTGCCATGTCTTTTAACTTGTTTTGAGCACAGACTAATCTTGTTTA
 TGCATGTTTTTTGATGAGAATAGGCTACTCATGAGAAATCTGTAAACCTAACACTAGTCC
 CTTGCATACTCTAAATTGTTGCTAGAATCTTAAAAATTTAGCACCAGACCGACCTTAGAA
 ATCATTAACTTTGGTGCTTTGTTCTACAATACAAGGAGATGGAATATTTTACCCAGGATT
 GCTTAGCAGGTTACAGTCTGCCCCTCTGAGTACCCAGCACTTCCCTGTGGGCAACATCAA
 CTTCTGATTTTCAAGTCTTAATTAGTACTCTGAAGAATCCTACTTGTTTTTAACTCCCA
 TTTGCTTTGAAGTGACTTTACCTGATTTTTTTAGATCCCTTATTGCAGCAATGCCACTAA
 GAAACTGAGTCTCTAGCTTCTTGGTGGGCAGGAGCTGCTTTGTGCTGCTCAGAATCATC
 CTTTTTCAGTAAGGGAGATATTGAAGAGAAATCTACTGAGGAGTCTGGGGGTGAGGCACTC
 AGGGAAATCCTGCTCCAGTCCACAAAAGCAGAGAGGAAGGGTTGGTTACCTAGAGTATTT
 AACATGCAGAGGCTTTGGATTTTACTCCTTTAATCCTTGGAAATGCCTATGGAAGGGGAA
 AGGAAGTAAGATGGTGACTCCAGCTTATAGACATACTAGTGTTACATATATTTAACTAT
 AATAGGAGGGTATTATTAGTTTTACTTAACTTTCAACTGTGAAGGATTATACTTCTCAAT
 ATTTGTCTCCAGTGCTATTTCAGTGATTTTTTCACTTTTCTTGAAGCAGCATGTCTGTT
 GCAAACTTCTAGAAATAATGAGAATATTTATATATTAGATCAAGCCATAACTTGATGAT
 ATAGTCATTTCTCTTATATTTTTTACTTACATTTTACATTTTAATGATTACTTTCATT
 TTTGAAAACATGTCATGCTGAGATGTATTTTTCTTCACTCTGTAATTAGTTATGAAACA

FIGURE 14.3

22/64

GTTTTTCTAAAATGCTGAGTATATCAAGTCTTGGCTAAGAATAAGTAATAAATATTTGC
 CACATGAAAGACTACACATATAGCCAGGTGCAGTGGCTTGCACCTGTTTTCCCAGCTACC
 CAGGAGGCTGAGGCAGGAGGATTGCTTGAGCCCAGGTTTCCAGGCTGCAGTGAACATG
 ATTGTACCACTCTACTCCAGAATGGGTGACAGAGCCAGGCCCCATCTCTCAAAACAGAAA
 AGAAAGATTACATAGACTACATATACACCCCCATCCAAAACATACACACACATCTACTTA
 ACCTAAAATGGTAAGAAGATAAATTCTTATTTTCTAATATATGACACAGAAAAGTTTTTT
 TAAAGTAGTTTTAAATTTTTAATTTTTTCTAGGTATTTCTCAAGCCATGTTCCCATGTGG
TATCTTGTCAACAAGTTGAGGTGGAACCCCTCTCAGCAGATGATTGGGAGATACTGGTAA
 AGAAAACCAAATAAGAAGTATCTCATTTAAGGTAAATTACTTCACAATATCAATGTCTT
 TAGCTTTCTCTAAGCTTTATTATATATTCTGAGTTGGTTTTGAATTATAAGAATGAATTG
 GGGCCAGGCACAGTAGCTCATGCCTATAGTCCCAGCACTTTGGGAGGCCAAGGCAGGTGG
 ATTGCTTGAGTCCAGGAGTTCAAGACCAGGCTGGGCAACATGGTGAAACCCCGTATCTAC
 TAAAAATACAAAATTAGCCAGGCATGGTAGTGCATGCCATTAGTCCCAGTCACTTGGGA
 GGCTGAGGCAGGAGAATCGCTTGAGCCCGTAAAGTCAAGGCTGCAGTGAGTCAGGATCTT
 GCCATTGTACTCCAGTCTGGAACAGAGTGAGACCTTGTCTCAAAATAAAAAAGAATGA
 ATTGATAGAGATCTAATGTACAACCTGACAACATATAGGTAATAAAATTGTATTGGGGATT
 CATGTTAAATGAGTAGATTTTTAACTACTCTTACCACAAAAACACAAAAGTGGGTAACGT
 GAGATGATGTATATGTTAATTTACTTCACTATAGTAACCATTATACTATCTATATGTAGC
 TCATAACACCATGTCGTGTATATTAATATGCACATTAAAATTTGTTTTTAAAAAAGA
 ATTGAGATTTTTTTAACTAGATATGGAGTGGACAAAATGTAAAGTGAATTGATCTTTTC
 GTCTGTGTGGTTCTAGGAGCTGCATGCTGTTTTCCCTTGAACAACATCTTCTAGATCAAAT
CGAATGATTTTTTCCAAAAGCCATTTTTCTCTGTTTGGGTTGATCAACAAACGTACATATTT
ATCCAAATTGGTAGGTGCTATTGTAATATTTGCTGTCTATTTCTACACTATAGCATTGAG
TCCAAAGTAGAAATGAATGTGCACTAATGAGCTTTATTTTCTACACAGTTGCACTAATAC
CAGCTGCCTCTTATGGAAGGCTGGAACCTGACACCAAACTCCTTATTCAGCCAAAGACAC
GCCGAGCCAAAGAGAATACATTTTTCAAAGCTGATGCTGAATATAAAAAAATTCATAGTT
ATGGAAGAGACCAGAAAGGAATGATGAAAGAACTTCAAACCAAGCAACTTCAGTCAAATA
CTGTGGGAATCACTGAATCTAATGAAAACGAGTCAGAGATTCCAGTTGACTCATCATCAG
TAGCAAGTTTATGGACTATGATAGGAAGCATTTTTTCTTTCAATCTGAGAAGAAACAAG
AGACATCTTGGGGTTTAACTGAAATCAATGCATTCAAAAATATGCAGTCAAAGGTTGTTT
CTCTAGACAATATTTTTAGAGTATGCAAATCTCAACCTCCTAGTATATATAACGCGTCAG
CAACCTCTGTTTTTCATAAACACTGTGCCATTCTATGTTTCCATGGGACCAGGAATATT
TTGATGTAGAGCCCAGCTTTACTGTGACATATGGAAGCTAGTTAAGCTACTTTCTCCAA
AGCAACAGCAAAGTAAAACAAAACAAAATGTGTTATCACCTGAAAAAGAGAAGCAGATGT
CAGAGCCACTAGATCAAAAAAAATTAGGTGAGATCATAATGAAGAAGATGAGAAGGCCT
GTGTGCTACAAGTAGTCTGGAATGGACTTGAAGAATTGAACAATGCCATCAAATATACCA
AAAATGTAGAAGTTCTCCATCTTGGGAAAGTCTGGGTTAGTATAAATTTTATAACTTGGG
 AGAAATTTTATGTGGCTTAAACATCCCCAAATTATGAATTAGAATAGTATTTTATATATA
 AATTGAAAATCAATTAAAAAGAAACACAGTGCCTAAAGGCCTTGGGGGACACATTTACG
 CTTTGCAGTAAAGTCCTTGTTTGGATAAAGATTGTATGTTTTCTGGCCAAGTAAGCTTGA
 ATAGGTACAAGCTTAGATAGGTTTCAAGCCAGAGAGGTCAAAATTACTTGCCTGAGATTGC
 ATAGCTAGTGTTACAACCTAGGATTCAAACCCAGGCAGATTGACTTGGGGGTTTATCAGGA
 TGGAGTGCCCTACAAAGCCTCCCATCTTTAATGCTTGCAGATTGTTTCCCCAGTTACCGA
 AAGCAACTTGTTAATATTAGGGAAAAGGGCCAGTGTAGGGAGAGATCCATGGCATGAGGT
 AACCTTCCTGCTGCATGTGGTGGCACCTGGATTGGAATGCATCCAGGAGCTGCTTACCCT
 GCCGGTGTCTGCTCTTTAATTTGTGTATAACGGAGAGGAAGTAGACAGGGCAACTAGTGC
 TCCAGCCCCCTCATCCTGGCCACAAATATTAATGCTACCTTTATATGACATAAGTCACTAG
 TCCATTTATTGGAACCTAAATTTGAACCACTGTAAAGTAAGACTTCATAGTGATAAAGAG
 AGGAACCTGTGTAGGAAAGAGAATAAAATAGAAAGAGAAGGTTGTCTCCTTTTGTAGATTT
 TTTTTTTTTCTCCAACAGTTTTTACCTGTGACCTTTATACAAATAACTGACAAAGCATTA
 TCTCTTTGGCCTACATCATTTTTCTTTTCTATTTTTTTTTTCCACAAGATGGAGTTTCACT
 CTTCTTGGCCAAGCTGGAGTGCAGTGGCATGATCTGGCTCACTGCAACCTCCGCCTCCCA

FIGURE 14.4

FIGURE 14.5

24/64

TTTTGGCTGCTGTGAATAATGCTACAGTGAACATTGGTGTACAAGTATCTGTTTGAGTTC
 CTCTTTTCAGCTCCTTTGGGATATACCTAGGAATTATGTTTAACTTTTGGAGAAGCTGAG
 AAATCTTTAATAAATGATAACACAAATACCTTATATTTGCCAATGCAAATATGAATATTTT
 TGGCTTTTAAGAGATTGATCATTTTGGCACGTGGTTGTAATTAATAAATAATGTCCCATG
 TTGTTTCAGTATTAATATTGTAGCCTAAAAGAGTGCTAGACTGTTTTACTTTTTACTCAG
 TTAATTCTTTGGATACTGGTAGAGTCAGGAAATGAGATATTGAACCTAAAGATCTTTGCA
 GGTGGGGTCCAGTGGCTCACACCTGTAATCCTAGCACTTTGGGAAGCTGAGGTGGGAGGA
 TTGCTTGAGGCCAAGAGTTTGGAGAATAGCCTGGGCAACATAGCAAGACCCCATCTCTACA
 AAAAAATTAAAAAATAAAGCCAGGCGTGGTAGCTCACGCCTGTTATCCCAACACTT
 CGGGAGGCTGAGATGGGTGGATCACTTGAGGTGAGGAGTTGGAGACCAGCCTGGCCAACA
 TGGTGAAACCCCATCTCTACTAAAAATACCAAATATCGGGGCGTGGTGCTAATCCTGT
 AATCTCAGCTACTCAGGAGGCTGAGGCAGGAGAACCACCTGAACTGAGGAGGTGGAAGTT
 GCAGTGAGCCTAGATCTCACCCTGCCTCCAGCCTGGGTAACAGAGCGAGACTCTATTT
 CAAAAAAGTAAAAATAAAATTAGACACATGTGGTGGCACATGCCTGTAGTCCTAGCTA
 CTCAGGAGGCTGACTGAAGTGGGAGGATCTCTTGAGCCCAGGAGTTCACACTGCAGTGA
 GCTATGATTGTGCCACTGCACTCCAGCCTAGGCAATATCTCAAAAAAATTTTTTTAAAT
 AGATTATTAGGCCAGACGTGGTGGCTCATGCCAGTAATCCCAGCACTTTGGAAGGCCAAG
 GCAGGCGGATCACCTGAGGCCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCC
 ATGTCTACCAAAAAATACAAAAATTAGCTGCAATGTCTATAATCCCAGCTACTTTGGGAGCC
 TGAGGCAAGCGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGTGGAGACTGCGCC
 ACTGCCTCCAGCCTGGGCGATACAGCGAGATTCTGTCTCAAAGAAAAAGGAATTTGTTT
 TCCTGTTCTTTATCGTAGAGGGAGGAAAGGGAGAATGGGGTTGGAATGGTTATTGAGTGAG
 CCAGTATTGGTAGATGTATCACTGGGCATAGAGAAAAGGAGCATTAAAAACTTTTCCGC
 CTAACAGATGTTTCTTCAGGCTACACTGCACTCATTGTGCTAACTGTAATGTCAAATCCC
 AGACCTGTGCCTATAGAACATGAACATCCTTCATTGGATTGTTTGGTTCAGGCTTACACT
 TTATTAGGAAGATCAGATGTTAAAAATAAGGGTGTAAAGTTAAGTTCAGATATGAGGATA
 ATTCATTACTATTCTTTTTCTGGCAGCCTAAAGACATAAGTGAAGAAGACATAAAAACT
GTATTTTATTATGCTACAGCAGTCTACTACCACCATGCTTCCTTTGGTAATATCAGAG
GAAGAATTTATTAAGCTGGAACTAAAGATGGTGAGTACATTTGTTATTTTGACTTTTTT
 TTCTATTTAAATAGTTGTACATTTTAAATTGTTCTTGCAACCTGTCATACCTGTGAACAG
 TATGTGAATAGTGAAATATAATTATGATAATTAAACAGTAGTTTTTATGTATTGAAAAAT
 ATCTTTGGCCGGGTGCAGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCA
 GGCGGATCACTTGAGGCCAGGAGTTTCGAGAGCAGCCTGCCAACATGGCGCAACCCTATCT
 ATACAAAAAATACAAAAATTAGCCTGACATAGTGGTGTATGCCTGTAGTCCCAGCTACT
 TGGGAGGCTGAGGCAGAAGGATCACTTGAGCCCAGGAGGTCTGTGTTCTTGCCTGCAC
 TCCAGCCTGGGCAGCAGAGTGAGACCCTGTTGGGGGGAAAAAAGTCTTTAACTT
 AAATAAATTTGACATTTAAATCTTAAATTATTTTATCTCTGTTTCACTACTAACTCTGC
 ATTTATTACTTTCTTTTAAATAGGACTGAAGGAATTTTCTCTGAGTATAGTTCATTCTTG
GGAAAAAGAAAAAGATAAAAAATTTTTCTGTTGAGTCCCAATTTGCTGCAGAAGACTAC
AATACAAGTAATAGCATGTTATTGAATATTTAATAAAATACTATTTGTTACATATGATTG
 ATAATAAAGTATGAAGTTCCTTGTAAACACCTTGCAATGTGAAGTGTATTAATAAACCTGCT
 AAGAGTAAGGAATAACTTGATTTAAAAATATTTTATTCTGTAATCTTTTAAATTATCTGT
 ACAAATTATTGACTTAACTTAAATTTAAAAATGAATGCCTTAGCACAAATTAAGTTCCAAG
 AATAGAGTTGATCATGTTAACTGGTAAATGGATCATGATTTAAATTTCTTCTAGGATTGA
 AACAAATGAAACGTAGTTTTAAGGGTTTGATTTTTTAAATTCCTATTTTTTACATGCAAT
 TTTACTGCACAACCCATCTTATTTGACAGTTCTTAAATTCGCAACTCTTCAGAAATATT
 ATCAGATCACTTTTCTTGTCTCCATAAGTTTTTTTTATTATTATATTATTTTTTTTTT
 TTTAAAAGACGGTGTCTCACTTTGTGCGCCAGGCTGGAGTGCAGTGGCATGATCATGGCT
 CACTGCAGCCTCGACCTCCCAGGCTCAGGTGATTCTCCACCTCAGCCTCCCAAGTAGCT
 GGGACCACAGGCGAATGCCATGATGCCTGGCTAATTTTTGTATGTTTGTAGAGATAGGG
 TTTCAACCATGTTGCCCAGAATTGTCTTGAACCTCGGGTTCAAGCAGTTGTTCTGCCTTG
 CCCACCCAAAGTTGTGGGATTACAAGTGTGAGCCACTGCGCCAGCTATTCTAGAAGTAT

FIGURE 14.6

25/64

TTTAAGAGTCATCTTTTTTTTTTTTTTTGAGATGGAGTCTCACTCTGTCAACCCAGGCTGGA
 GTGCAGTGGCACACTCTCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATTCTCC
 TGCCTCAGCTTCCCTAGTAGCTAGGATTACAGGCGCATGCCACCATGCCCTGCTATTTTT
 TGTAAGTTTGTAGTAGAGACGAGATTTTACCATTGTTGGCCAGGCTGCTCTTGAACCTCTGAC
 CTCAAGTGATCTGCCCTCCTCAGCCTCCCAAAGTGCTGGGATTCTAAGTGTAACCACCA
 CACCCAGCCAAGAGTGGTCTTTTTTACAATATTATTTTTTTGATTAGGACATTCACTCTTGT
 CATAAAATTGAAGATACTCTAGTCATTTAGAAATTTCACTGTTTTGGAAGTAGACATTGTT
 TCTTTATTTTTGAAATGTTATTGAAGGAATACCATTGAGAGATAACAAATGTAAGAAT
 TGTGAAAAGGATAATTGTGACACAAATCAAAATTATAGATAAAAAATATACCTGTAAATG
 TATTAAGGCAATAACATTCTTCTGCTTGTGACCATAAATATTTATATCCCTGGATGG
 GTACATTGTTATTGTCAAGGGTGTAAATAATGATCTTGCATGCATAATTTATCTCTC
 TGGTATAACAGAATCAGCAATTTAGTTTTCTGGGACCCGAGAAAAACATGCAAAAGACAT
 ACTTTGAAATGTAAACTGATTTTTCTTGCACCTGTAGTCTCTCTAGATCCTATGGTA
AAAGAAGAAAACAGTGAGGAAATGACTTTATCTTCTTTTTTAAAGCTGAGCTCTTTG
GGGTAAGAAGTTATGGCCAAACTAGCATGTTAGACATGTTTTTAACACTATATCTGGCAG
 AGTTTTCAATGTAAATATTAAAGTAGATGTTAATGTCAATAAGTGATCTTAATAATGCAT
 CAGTAGATATTTTTTCAAGGATTGTCTCTATCTTCACGCCTAGCTTATAATTTGCCTTGT
 CGTCTTTTTTTTTTCTCTTTATTTTTATGTTTTATCCATCCCTGGTGGTAGGGGATAA
 CTTGTCTTCTTCGATAACAAGAAGTCTGAAGCTTATTAGAAATTTACTTTGAGAATTG
 ATCGATGAGAAGAAAGCAACTAGATATCACGTGGATCATATATGCTTGAATAAAACAATA
 ATTCTTAGAACAAATAAATACATTTTAAAGTTAAAGCCAAAAACATTAGTTGAATGTTT
 AAAAAATTTCAAATTAAGTTATTCCTTCACTGTCTTGTATTACTGTAATAATTTGGATT
 ATTTGTGTTTTTCTCACTTTTAAACAAATATTAAATAATTTCTCTTTGATTAAAGTA
 GGGCTAGATAAAATATAAAAAATATTTTTTAACTCCTCTTAATTTCCATATTTCTTATA
 TAATATGAGAATCTCTTATAAAACACTACCTCTTAGAAGTCTCCACAGAAGCTTTGGTAGA
 TGTAAGTAGTAGGGATTTGATTTCTTAGAATGGTATAATCTGTAAATGTTTTAGTAAAGG
 ATTAACGATAAAGTCAAAATGTTTATAGCACAGTGTATTAATATAAAATAAAATCTC
 TTTTTTTTTTTTTGAGATGGACTCTCACTTTGTCACTCAGGCTGGAGTGCAGTGTGCAA
 TCTCAGCTCATTGCAACCTCCGCCTCCTGGGTTCAAGCAATCCTCCGCATCAGCCTCCT
 AAGTAGCTGGGATTACAAGCATGCACCACCACACCTGCCTAATTTTTTGTATTTTAGTA
 GAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTCAAGTGATCCGCCTGCCTCAGCCTC
 CCAAAGTGCTGGGATTACAGGCGTGAACCACTGTGCCAGCATAAAGTAAATCTCTCA
 GACTCTCATGTGATCATGTAAAGTGGCAGGCAGTCACAGTCAAGAAGTAGTTTAAAGTTC
 ATGTTTGTAAATATAATCTACAGATTGATACTGGATTTTATAGGTAATGTTTAAAGAGAA
 AATAAGTTTTTAGTTATCTCAGTACTTCAAAAGCACCCATTTATGATTATGTTGATTAC
 TAACTAAATCATTTGGGGGCTAGAGGTGTTTTTTATGTGTTAAGATTCTTAAAGGAGT
 TCTATTAGGGCAAACTTTTAGTAAGTGCATATTTTAAAGTAATAAACTAATTTTAA
 AGCTTGGAGGCTGGGCGCGGTGGCTCACACCTGTAATTCAGCACTTTGGGAGGCCAAGG
 CGGGTGGATCACTTGAGGTCAGGAGTTTGAGACGAGCCTGAGCAACATGGTGAAACCTTG
 TCTCTACTAAAAATACAGAAATTAGCCAGGTGTTGGTGGTGGGCACCTGTAATCCCAGCTA
 CTCGGGAGGCTAAGGCAGGAGAATTGCTCGAACTTGGGAGGCAGAGGTTGCAGTGAGCCG
 AGATCATGCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCAAAAAAAAAA
 AAAAAAAAAAGCTTGAAGTCAGATTCGACATTAATCAGTATACTTTCTCTCAAGTAGGGG
 ACAATTTCTAAGATTTTAGTCTTTTTAAATTTATTAAGTCTGAGCATGGTGGCTTGT
 GTCTATAATCCAGCACTTTGTGGGGCCGAGGCAGATGGATCACTTGAGCCAGGAGTTG
 GAGACTAGCCTGGGCAACATGGCAAAACCCCGTCTCTACAACAAATGCACACACAAAAA
 CCCAATCAGCTGGGTGTGGTGTACACTCCTGAAGTCCAGCTACTCGGGAGGCTGAGGC
 AGGAGGATCACCTTTGCCAGGGCGTTTGAGGCTGCAGGGAGCTGGGTTCACACCACTGCG
 CTCCAGCCTGGATGACACAGCAAGCCCCTTTCTCAAAAAAAAAAAGATAAAAAATTAAT
 TAAATTAATTAAGTACACTGGGAAGGCAAAATTCAGCATTTTTTATAGCTAAATTTTAT
 CCTGCTTCAGTCTTTTATCATGTAAGTATGTATTTTTTACAGAGGAGTGAATTCCTTA
GGCGTATCCTCCTTGGAGCACATCACTCACAGCCTCCTGGGACGCCCTTTGTCTCGGCAG

FIGURE 14.7

26/64

CTGATGTCTCTTGTGTCAGGACTTAGGAATGGAGCTCTTTTACTCACAGGAGGAAAGGTA
AGTGGTTAAGGTGTGTTCAATTTTCTGTAACTTAATAACTTTTCATTTATCTTTCTTT
GGGTTTTGACCATCTATTATATAGGGTGGGTTTTGACCATCTATTATATAGGGTTTATAC
GACATATGGAAAGCATTCAATTTATTTACTAATATTTCTGTGTGTCTGCTTTTAGGTGTTG
GGGGAGTGATGACGAATAAGACTGATGTTCTCCATGCCCTTTTCTGTGTCAAGTTGATAC
AATTATATGGTTTTTCTTTTTTAGGCTATTAGGTGTTGATAGGGTTGAGTAACCTTACAAA
TGTTGAACCAGCCTTGCATACCTGTGATAAATAACCACGTAGTTGTGGTGTATCATTCTTT
CTACATTGCTGAGTTTTATCTGCTAATGTTCTGTTGAGCTTTTGTCCATTTAAGTTTGAA
AGTGATTAGTTTGAGTTTTCTGTTTTTGTGTTGTCTTTGTCTGGTTTTGCTATCCGTGT
AAATCTGGCCTCATAAATGAGATGGGAAGTATTCTCTCCTCTCTTTTGTTTTTTTTGGGA
AGAGGTTGTATAAATTTGAGGCTGAATCTTGGTGGTTGCCACAATGACAGGAACTATTTT
TGTGACTGAATATATTGGGAATTTCTATAAAGCAATTATTTTCTAGGGAAGTGGAATC
AACTTTAGCCAAAGCAATCTGTAAAGAAGCATTGACAACTGGATGCCCATGTGGAGAG
AGTTGACTGTAAAGCTTTACGAGGTATGAGTATGGTAACACTCTATATAAATCCCTTTTT
CATTAGAAAGACAGGAATGTTATACATAATGCTGTCAATCTAATAAATACACATATCATC
TAGTCTTTAACTTTTTCTGTTTATCATTAGTCATTAAATTTCTTTGGCTTTCTAATGTT
TTTGATAAAATTTCTAAACTCTCCATATTTAATGGAGGCCTATTTTTTTTTCTAGCCAG
AACTTTTTGTAGACTACATTTCTGGAAGTGCTCACTGACACCACTCTGAAAAATTAGTAC
TTAGAATATACTCTAATTGGTATAAATGATCTCTGAATTGCTATGGAAAACTGGGAGAAT
GGTTGCTTCAGGGGAGAGAAAGTAGGAGGCTGTGGACAGCAATGAGGAGAATTACAGTTC
ACCATATAACACTTTTGTACTTTTAAAGTCCTTAACATTTACATTATTATCTATTCAATT
AAAAAATATTGGGAAGATTTTACTTTGAACAGTTAATTTTTCCCCCATGGGTACCGCTGT
CATATAGTTCCAACTAATCATGAACCTGTGTATTTCTGTTCTTTGTAAATTTAACTTT
GTAACCTACCAGGAAGTTTGAAGCCAAATTTGTGTTTCAAATATAGCAACTCCAGGATCT
CTAGGCAGATGCATTTGCATTTGATTTTAAATGAATCTTGATCCCTTACTCTCACTTATG
TTTTCCCATCTCTACTTTTTTTTATTTTGTGTAAGCCATCTAAAATTTCTCAATGGGATG
AACTGGGTATAAATGAATACATGCATACAGGAATTATAGTAGCATATTCCTTTTCTTTT
TTCTTTTTTTTTTTTTTTTGGAGACAGAGTCTTGCTCTGTAGCCCAGGCTGGAGTGCAGTGG
TGCGATCTCGGCTCACTATAGCCTCCACCTCCCAGGTTCAAGCAATTCTCGTGCCCTAAC
CTCCCAGTAATTGGGACTACAGGTGCATGCCACCACACCTGGCTAATTTTTGTATTTTT
TAGTAGAGATGGGTTTTACCATGTTGGCCAGGCTGATCTCAAACCTCTGACCTCAAAGT
GATCTGCCTGCCTTGGTTTTCCCAAAGTGCTGGGATTACTAGCATAAGCCACTGCACCTGG
CCTCCTTTTCTGAGTTTTATAAATTTGATACTTTACTGCACGCTTTGAGACTGTATTAA
TTGAACCATGTTGATGAACAAGTTTTTGTGATGGGTATATTAATAAATATAGATCAAAAT
TTTTATAGTTAAATCAATATCGAGCTTTTCTAGTGCTTTCAAAGGACAACCTGAATTTT
CCCAGCACTGAAATGATACTGAAACCATTTTCATATCTTCTGTATTAAGGAAAAAGGCTTG
AAAACATACAAAAACCCTAGAGGTGGCTTTCTCAGAGGCAGTGTGGATGCAGCCATCTG
TTGTCCTGCTGGATGACCTTGACCTCATTGCTGGACTGCCTGCTGTCCCGGAACATGAGC
ACAGTCCTGATGCGGTGCAGAGCCAGCGGCTTGCTCATGGTAAATGCATCCACCACTGGC
TTAAGGTCTTGTTCTTTTGTGTCAGTCAGCATTTTGTAGTCTTAACAATAAATCTACTCTCT
CAGAGAATAATATATGTGTTATGTTAAGTGTGTTGAGGCCCCCTGATGGCATTCTAC
AGTTGTCCTATAGACTGTAATAGCAAAATTTGGTAGAGTAAAAACAGTGTGAAAAATCTGC
AACTTCATGGTTAGTCCTTTAGGGTTTTTCACTCTCCCTTACTTATTGTTTAAATTTACAG
ATTTACTCTTTTGTTCATTTGACAAATATTTGTCAAATGCTTGTGCACAGTCTGTATTCT
CAAATTTCTAGGAGAAAAAGAAGGGTGAACAGTATTAGCGCAGAACGATACTAATAATGAT
GGCTACTGTGTATGAGTAGCCAGCCCTTTCTTGCTTTCTTGGATTGCTTTGTATTCTAC
ATGAAGATATTCCCTGGGCTTTACAGGTCAATAAATGGAAATTCAGAGAGATTAATTTGA
CCAGGGTGACCAACAAGGAGATGACAGCATACACTATGCGAGAAGTATACACAGAGTAGT
GTAGGAGCATATAACCTAACTGGGGGTGAGGTGGGATAAGGAGTTATCAGGGAAGGCTT
TTTGGAGGAGTTGACAACCTGAGCCGAGTTTTGATGGAAGAGTAGAAATTAGCATGAACCA
ATTTTCATGCTAATAAAGAAGCAAAGGAAGCGTGGTCTACAGGCAAAAGCACAGAGGTACA
GGAAGTAATGATATGTTGGGAATACCCTGTTGACTGGAGCTTAGAGTGCAAGGAGAGGA

FIGURE 14.8

27/64

GTGCTAGGGAGGTGAGGTTGGAGGGTTTGGCAGCATTGACTTGCTTCAAGGTTCTTAAGA
 GCTGAAATAGATATAAAATGCAACTAAGAGTGGCTTGGATTATTATTACCTAGTGTGTTA
 ATCTCAAATTTTGAAATCTATAGCATCTATAGGACTGGTGTACTAATCTTACACTCGAT
 CTGTTACTGTTCTTATACTAGATCTATTAGTCCAGTGTTTAAGGGAGTGGTGCAGATTTC
 TAGGTGAGGACAGGACTCAGATGTACATTATTAATGCCTATTTTCAGTTCTGACCTTCTCA
 TATGAAACCTTATAAGACCTGGGGTAGGAAGAGATTGTTCTGGAAGTCATAGGAATATGA
 ACTGTATTTTGTTTAACAACAATACAGTATGGAAATTTATCACCTTCCAGAATATTTTA
 TTTTCAGAGACAAATTTTATCATTCGTTTCATTTATTTTCATAAGATCCACGAGTAGGGAAC
 CTCCTAGACATTGCTCTGAGTATATGGTCTGAGTTTGCAGTACCTCTTGTGTCTCCATT
 AGATTTTATTAGGTCTCAATAGATAAAATCAGGGAATAACTAGATGGATTCAATTTTTTAAA
 GACATGAAAGAGCGATACCATACTACTGCACCTTAAAGGTCAACCTTAGAGTATCATTAA
 TTTTAAATGAATGTATAATTTTAAATTTTCATGTTTACTTTTCTTAAGCTTTTGCACAT
 ATTGCTTAATTCAGCTTTGAATGATATGATAAAAGAGTTTATCTCCATGGGAAGTTTGG
TTGCACTGATTGCCACAAGTCAGTCTCAGCAATCTCTACATCCTTACTTGTCTGCTC
AAGGAGTTACATATTTTCAGTGGCTCCAACACATTCAGCTCCTAATCAGGTAATACACT
 ACTTGTAAAGGATTATTGAATTATGTCCCTTTTATAGAAATTATTTTCAATTTTATTAGT
 AATTCGTGGCTTTAAATTTATGCTTCTCTTAATGATTTTAAAGGATATGTAAGTCAACATT
 TGGTGCATATTGTGCTAGAGGCATAAATTATAATTTATAGCCACCTGAAATGTTAGTATG
 CGCTTTCCAAGAAAATGACTTTTTTGAAGTGGTATTTCTTTGAATGAGAAAGAACAGAG
 AGAAATAGATAGATGGCTTTTAAACACTTCATTAATTAAGTCTTTTTCACCATCAC
 ATAATGGCACTTAGTCCCTTTTGGGAACCTCATGAGGGTTTGTAGTGTAGTGTGAGCTGAAAG
 AAATATGTTCCAGGACTGGCAACATATTCTAAATCTTTAAATTTTACCTAGCATCT
 ACCCTAAATATTAGACCTGTGCTAGTTAACTGCTATTGAAGAACAAGGTATTATATC
 TATTATTAAGGATAATAGAATGGTATTTGAGATATTGGTCAATTGAATATGAATATGTTTT
 GAGAAATAAGTTTTATAGGAACCAAAAAAATTTCTTAAAGGAACCATATATTACTAAAA
 ATGCTTCTTATTGGAGAAAGAAATGACAATCATTTATTAATGTGATTTTTTCACACTTT
 ATTAAGATATAATTTAAGTACAACAACTCACATAAAGTGTACAATTTGATCAGTTTTAA
 CATATGTAGATGCCATGAAACCATCACCAATTAAGGAAACAAACATTTTCATCACTCC
 AGAAGTCTCCTAGCCCTTTTACTACCCATTCTCCCTGCTCCATCCCAGACAACTACC
 AATTTGCTTTCTGTCACTATAGATTTGTCAACCTGATTTTCTCCAAATATACATTCAAAA
 ATATACAGTTGAATACAATTGGAAATTCGAATTTTGTGTTTTTTCTTTAGGAACAAAGA
TGTGAAATTCGTGTGAATGTAATAAAAAATAAATTGGACTGTGATATAAACAAGTTCACC
GATCTTGACCTGCAGCATGTAGCTAAAGAACTGGCGGGTTTGTGGCTAGAGATTTTACA
GTACTTGTGGATCGAGCCATACATTCTCGACTCTCTCGTCAGAGTATATCCACCAGAGAA
 AGTATGTTTTACTATTAAACCTGAACTTGGAAATCTTCTTTCTATTGTGGAGAAATGTAA
 TTGTAGTAAGACAAGAATTAATATATTCCATTGTAGTATTTGAATAAGCAGTTATTTGA
 GTAGAAAATTAGTGTTTCCAGCTAAGATGATGGCATATTTTGAAGTTTCATATAGTGAAT
 ATAAGTAGTAAAGAGTTTGTATTATTTTAAACAGAAATTAGTTTAAACAACATTGGAC
TTCCAAAAGGCTCTCCGCGGATTTCTTCTGCGTCTTTGCGAAGTGTCAACCTGCATAAA
CCTAGAGACCTGGGTTGGGACAAGATTGGTGGGTTACATGAAGTTAGGCAGATACTCATG
GATACTATCCAGTTACCTGCCAAGGTATGTTTTAAAAAAGAAAAAGTGAATACTTACTCC
 CAGAAGAACCCTGTATTATTGGCTTTGGCTTTATGTGTGCTGAGCTTGCCCAATCTCCGTGT
 GAGTCAACAAGTGTTTACTGAGTTACCAATAAATGTCTTAACACTATTTTAGGTACTTT
 AACAAATTTTAAATTTTATTAATTAATTTTTTATTAGAATTGAGACCTCACTCTGTCTCT
 AGGCTGGAGTACACTCACAGCTCACTGCAACCTCAAACCTCTGGGCTCAAGCAATCCTCC
 TGCTCAGCTCCCCAGTAGCTAGAACTACAGGCATGAACCACCATGCCCGGCCAATCT
 TTAATTTCTTAGAGACGGAGTCTTGCTATGTTGCCAGGCAGACAGATTTTAAATGTGTA
 TGATGCAGTCTTTGATGATAAGAACTTATAATGGAAAGCTGAGGTGATAGTTACAGTAA
 ATACATTTTATGATGATAAATCTGTTTGTCTTAAATCATTCAAATTGTAGTAAAGCAAGATG
 AACTGTCTGCTGGGATTTGAGCAGAAATGGATAGGAATAAACTAGGAGGTAGAAGAGTTA
 TCAAGGTTTACAGGACTGATGGGTGAAGCTAGATTTCCAGACCCGGGATGTCAGTCCCTG
 AAAAGCAGACTTGGCAGGCATAGACGAGGCAGATAGCAGGATAAAGGAGACAAATGTAGA

FIGURE 14.9

28/64

TTGTTCTTCAGAAGATCAGATGGTAGAGTCTAGGAGGTAGTGTGTTTTAATCAGAGATCT
 GAGAGGCCAAAGATCATTGCATGAGATCAGGGACCCATGCAAAGGAGTGAGAAAAAACT
 GGGTTAAGGAGCCTGCTGCATGGCAACTCCTGGGAACAGTGGCCACTGGGGCCTGGGACA
 TGTTGATTGCAGCCCAGGACTGTTAAAACCACTGTGAGAGAACATGGGTATGGAAGTACT
 AGCTAGCAGGATCATGACCCCGATGCTGGGATGGGGCATCAAGCATTAGTACATGGAGAT
 TCAGTACATCCAGATGCAGTACATGGAGACTATATGCGTAACTGCTGACTTTGGGCCTCT
 TTCAGATTGGAGCAGAGGTAGAGGTGAGTGGGAATATTCTCAATAGAGGGAACATAATAG
 GCATACCTAATAAAGGAGACCAGGATATTGCAGACAGTAGCCTCATGTTTGGCTCACCTG
 TTCAAAAAGTTCTCTTGTCTTTGAGCAGTGGTGCCTTAAAAGGTAACCTTGAGAAGCAGTC
 GATTATTTGTTTCAGCCTGGAGACTCTTGGGATATTTTACTATCTTTGATTGAATAGATTT
 AAATGTACACAGCTCTCATAACTTGCCCCATGAAGCATATCCATGAAAGGCACTATACTT
 GTTAAAAGATTGGTTTTGTACTTTTTTAAATGTAGTACTTTTTAATAAAACAGGAAAAATAGA
 AGTTCGTATGCAGTTATATGCATTTTATATAGAATGTGTTCTTAATTGGAAAAAATTTGT
 CGTAGTTCCTTTGAGTTTCAATTTACAGTTTTTGTAGTAGGAATTGTATTTTCTACTGTTGTAC
 TTGCTGTTACTAAAGAAAGATGGTTCGTGATTACCATCTGAATTTTTTTTTCTATACATTGA
 TCTTTAGCTGCTACTTAGTCATTTCTGTTTAGACTTGAGCTCTTTTTCTATATTTTTTTTT
 TTTGTTTTCTCAGTATCCAGAATTATTTGCAAACCTTGCCCATACGACAAAGAACAGGAATA
CTGTTGTATGGTCCGCCTGGAAACAGGAAAAACCTTACTAGCTGGGGTAATTGCACGAGAG
AGTAGAATGAATTTTATAAGTGTCAAGGTATGTTGTCTACTTATCTTCTTTTTTTATTTA
 GGTAAAATTAACATAAATGCAGTTAGCCATTTCAAAGTGTAATTCAGTGGCATTTAGTG
 CATTCACAATGCTATGCAACCACCACCTCTCTCTAATTTCAAACCTTTTTTATTCCACTC
 CTCCTCTTGCTTATCCCCTGGCAACCATTCTGCTTTTTGTCTCTATGGATTGTCCTT
 TTCTGTATATTTTATATAAAACAAATCATGCAATATGTGACCTTTTTTGTCTGGCTTCTT
 TCACTTATGTAATGTTTTTATGGTTTATCCAGGTAGTAGCATGTATCAGTACTTCATTCC
 TTGTCATGACTGAATAATGTTACCATACTTTGTTTATCCACTTATCAGTGGTGAACATTT
 GAATTGTTTCTACCTTTTGGACTATTATGAATAATGTTGCTGTAAATATTTCATGCACAAAT
 TTCTCCACGGATATGTTTTCTCTTGGGTATAAACTGAGGAGTAGAATCTTGGGT
 CTTAGGGTAATCTCTAACTTTTCAAAGAACCACCAAACCTGTCTTTCACACCAACTGCAC
 CATTCCCACTAGCAGTGTGGGGGGTTCTGATCTCCACATCTTTACCAACACCATTATG
 TTTCTCAATTGTGGGCTAGTCTCACATTTGGAAAGCTAGTGGGAGCAGCGATCCATCTAT
 TAAAAGTTGTATGAAATTGAGTAATGAGCCACCTCTCTCTGTAGGGCTTATTATGTTCT
 TGCTTAAGGCAATCTTCATGCATTGTGAACAGAATTATACATAAATGCTCAGATAAAAGG
 GCAAACCATCTTAAAGGGAGTAGACAACCTAGAGGCAGGAGACCATACTGAGGCAGGAAG
 CTGGGGTTTTTATGTTTCTGTTACTTTTTGACTATATCTCACCATTGCTTTTGTCAAAGTG
 AGACTAGTCTAAGTTTTTTTTTTCAGGTATAAGGTGAGTGTGGTAATTAAGGGGCATGCTAG
 CAGATCATTTTGGGTAAATGCTTCACAGTCCACCCTGGTGTGTCTATTGTGGTTCGAGATC
 CAGTATCTTAGCTGTGTAATTTTACAGACATCAGCAATATTAGTTTAAACAAAGGGCAATTAG
 ATTCCAAGACAAAGGAATCGTGTATTATCTAGCCTTATTCAAACCTTGATTTATAAATCA
 GTTTAGTAATTTATTTATTTGTTTCTGTATTTATTTTTATTTCTTTGAGATGGAGTCTCA
 CTCTATTGGCCAGGCTGGAGTGTAGTGATGCAATCTTGGCTTACTGCAACCTCTGCCTCC
 TGGGTTCAAGCTATTCTCTGCTCAGCCTCCCGAGTAGCTGGGATTACAGGCTAATTTT
 TGTATTTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGAC
 CTCGAGTGATCTGCCCCGCTTGGCCTCCCAAAGTTCTGGGATTACAGACGTGAGCTACCG
 TGCCCAGCTCAGTTTGTAGTAATGTATAACTGGGTTTTTACCCAGTTGTAAATTACTCTTTTG
 TCGTGTTTTTTTGAGAACTGGCAATGACGGAGAACTAAAAGTGCCAGGCTGTTGCCTTG
 TTCTGTATTTTTGCTTATGTTTTTTTTTTTTTTTTTTTTTTTTTCTCTGAGACTGAGTCTTG
 TTGTGTTACCAGGCTAGAGTGGAGTGGCATGATCTCGGCTCACTGCAACCTCTGCCTCCT
 GGGTTCAAGTGATTCTGCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCCTGCCACC
 GCACCCGGTGAATTTTTGTATTTTGTAGTAGAGACGGGATTTTACCATTGTTGGCCAGGCTG
 GCCTCGACCTCCTGACCTCATGATCCACCAGCTTCGGCCTCCCAAAGTGCTGGGATTACA
 GGCGAGAACCAACCGTGCCCGGTCTTGCCTTAGTTATTTCTTGTTCCTCCTCTAGTCCTA
 TAGTTCCTGACTGTATTGAGGAAATGTAATTAAATATTATTATGTTAATAGATATTTAT

FIGURE 14.10

29/64

GTGGTTGAATATTAGAAATTCCTTATTTTTGGTCACATATCCTGATCAGTAGTTGGTCTTC
 TGGAGATAGTGATTTTTCTACTAGAGATGACTTTAGGACCTATTCAGGTTTTTTTAAAGAT
 CCCAATTTAAGGAAAGACTATTCTCATTATTGATTTTGCTATATGCAGGGAAATTTATTT
 CGAAAGGTTTTTCAGTTGGCTTTTAGGGAAGATTATATATTCTCTTTTTTTTTTTTGGC
 CTTTTCCCATGTTCTAAAAATGATATATTCTTTAACTCCTATGAAAAATACATTGTTTC
 AGTAATTGAAGATGCTGATTAAAGTCATATCTCTACACATTTTTTAAAAATTTGAGATAGA
 TGGGACTTTGTCCCTTCTTACACCATTCACTTATTCACTTGGAAAAACTATTATCCAATA
 CTTATGTGGCAGACACTGTTTCTGGCACAAGGGATTCAAGCAGTGAACAAAACCTGCCTTTT
 TGGAGTTTACATTCTACTAGTGGAAAGCGACAACAAGCAGATAGACACATTCAAGTATATA
 ATTCAGTGTGAGATGGTGGTGGTAAGTCCTATGTAGGAAGAAAAGCAGGGTAAGGAGGCT
 TGGAGTAACTGGAGTGAGTCATAGATGGACTTGTGAGGAAAGGGTTTCTGAAGAGGTGGT
 ATTTGGGCAGAGATCTAAATAAAATGAAGCAACAAGCCATGAGAATATCCGGGGGAAAAT
 GTTCTGGGCAGAAGCATCAAGCATAGAAGTGTGGTATGATATTTATTCTAGCACACATT
 AATTTTAAAAATGTATAAAAGACATCCATTTAATCATATTAAAGATTTCCATGATTCAAT
 TAGACTTAGTTCAGAAACCAATTTATATTTTTCTTTTTAAATAATTTTATCTCAACTCTTA
 TTTTACCCAATAGGGGCCAGAGTTACTCAGCAAATACATTGGAGCAAGTGAACAAGCTGT
TCCGGATATTTTTATTAGGTTGGTAGCCTATGAATGTTTTTAAAGTAACTGACTCTGTTA
 TTATTTATCAATCAGTGCTTTTTTTGGTCTTGTTTTTTGAAGAACTGATATTTGAAACCT
 GTGGTTTATGTGAATTATTAATAAGCTAGAGGACGTGGATTCTCTATTTTCATCAAATAAT
 ACAAACATTTTAGATATTAAATTTTGGAAATTATTTGGTTTTGTTTTACAATAGAAATA
 CTCCTCAAAGTGGAAATCGAAGTGGTTATTCAAAGAAATCTCAGAGTAGATTCTTATATGA
 AGCAAATAATGCCCCTAATTTATCTCTAAATTTGTAAAGTTCTAAATTCTTTTTTCCCC
 CAGTTTCTAATTTATCTTCTTATAAGTCAAGAGTCCATCTGGCCAATTTAATTTTCAGTGAG
 TGTAACATTTTTGCATATATTAAAAAACTGTATATGAATACAGAAGATGGTATTTAAGGA
 TGAAAAATAATTATTCAAATGTGATAGCATTATGGGGAGTTTTTAAAAATAAAAGTTACTGTT
 TTATTTCTTCAAAAAATTTTATTATAAAGTATACAGTTAAGAGAATATACATAAAATACAT
 ATGCAGCTTAAGGAAGAATAATAAAATGAATACTTCATGTATTCAACACCGAGTTTACCA
 GGAAAAAGCATAAACAAAATAAACCTCTTCCACGTAATTCCTGGGTAAAGAGAAGTTAT
 AGTGGAATAATTTGGGAGCAAACGATAATGAAAATACTATCCATTAAAATTGTTAGATG
 TTGCAAAACTGATTTCAAGGAAAATTTATAGTGTTAAATGTTTAGAAAAGAAAAAGGTT
 AGAAGTTAACCCTTATGTATCTATCTCATGAAATTAGGAAAATTATAGATATAAACTAA
 AAAATATGTTAAAGGGGAAATAATAAAGATAAGAATGAAGTTTAATGAAACACAAAACAG
 AGAAGCTCACAAAGCCAAGATTTATTTTTTTGAACACCGAGTACAATTGACAAATCTCTAA
 CAAGTTTGATTAAGAAAAAAGAAAGCATGAATAAACAATTTTAGGGATAAAAAGGGAAAC
 ATCGCTAAAGATATCCCAGAAATGTAAAGATAATAAGGGAATATTATGAAAATATTCAT
 GCCAATACATTTGAAAACCTTAGGTGACATAGACAAAAACAAAATTGACCAAAATTGAGCA
 AAAAAGAAACAAAATCTGAGTAGTCTGTAACTTAGTAAAAATTGAGTTAGAAAAGTTAA
 AGAAGTCTTTACACAAATCAAACATCAGACTCAGTTTTCTAGGAGAGTTTTGCCAAACAT
 TCAAGTAGCAGATAATTCTGGTCTATTTTTTGGCCCCAGAAGATATATTTTACTTGCCATG
 CATTTAATGAGATAGCTGTTGATTTTTTTCAATCACCGTGACAGGTGTTTTATATTAGGT
 GTTATTCGCCAGACATCTAGTCCACCTGTTGCCAGATATGGAATTAATATTCACTTATTT
 TGAATTTAAATTTGTAAATAAATTAATAAAACAAAGTCAAAGTTCAAATTATTAATAAAG
 TAAAAGAAATAAAATATATTTTTATAGAGAGCCCTTACAAAACAGTACCAACATAATGAGC
 TTTCCAAATTTTGAATGGGCAAAATAAATGAATAGGCATTTTCAAAAAGAAGGAAGGGTG
 GCCAATAAGTATATATTAATATAAAAAATGGTTACTTGTAATAGGAATCAAAGTGTTTGA
 CTTATTGACTAAGAGTCAGTTTTTTGTTTTGATCCCTGTTAGTCTATCCAGAAGGCATGGG
 TCTTAATAAACACCTTGACCTCAACAGTTTACTGAATACAAGGGTAATTTTCATATGCCTT
 GCCTTCTTTAAGGGTTTGTGTAAAGATTAATAAATAACATAAAATATATATAAATACAT
 TTATATGTATTTATATGTAATTACATACAACTTGCCTTCTTTAAGGGTTTGTGTAAAAA
 TTAAGAAGATATATAAATATATATAAATAACATAAAATAAATACATTCAATATATGTATAT
 GAAATCACTTTGCCAATATGAAGCCTGATTCAAATATGAAATGTTGTTTGTTCCTTCCCA
GAGCACAGGCTGCAAAGCCCTGCATTCTTTCTTTGATGAATTTGAATCCATTGCTCCTC

FIGURE 14.11

30/64

GGCGGGGTCATGATAATACAGGAGTTACAGACCGAGTAGTTAACCAGTTGCTGACTCAGT
TGGATGGAGTAGAAGGCTTACAGGGTAATAATTATAAATACAGAAATAGAATGTTATAAC
AAAATGTCATCATGTATCAGATTTTGGTAAAAAATGTTCTTTTCTCTAGGTGTTT
ATGTATTGGCTGCTACTAGTCGCCCTGACTTGATTGACCTGCCCTGCTTAGGCCTGGTC
GACTAGATAAATGTGTATACTGTCTCTCTCTGATCAGGTGACAATTCATATTTAGAGT
CCAAAACCCAACAATGCTACACTCTTTCTTTGTGAGCTTTACTTCTGCCAGGTAATGGC
AATTGTCCTTAGAAGACCAGCTTTCTTAGGGAAAAGCTTTAGCCACTGTTTGCTCAAAGC
ATAAAAAGATTCTGAATTAGATGCAAAGCCTTTTTTTGGCCAGTGCAAGTCTGAAAAC
TTGTAATCCTTCTGTGTTGGCTGATTGGGGAAAAAATGCAAGAAACCTAATGTATTA
TATTTTACATTATCTTCTGTTCAAAGATTACATACTTCCATTATCTGTCAAAAAA
ACTCTGATACAGAATCAAGCATGTGAATCGTAAGCATGTAAGCAGGTTTCATAGAGATAA
TTTTTCAACTCTTCTTGTCTGTGTTGTTCCAACCTTATTCTCCAATTTAGAAGCAAA
CAAATAAATGAATGAAAGAACAGATAGACAAATGAATAGTCAAAGGTATAAAGTATCTGT
ATATATGTTACATGTAGCTATTATTTAAATTATTTAGATTTTCTTTTGAAATACCTTCT
TGGCACACTTGCCTAAATCTAGAAAATAAGCACTGTGTGAATAAGAAATTATTTACACTG
AATATTTTGTAGGTTTTTGGGTTTTTGTTTTTTTCAGACAAGGTCTCACTTTGTCACCCAGG
CTGGAGTACACTGGTACGATCACAACTCACTGCAGCCTCTATGGCCCAGGCTCAAGCAAT
CTCCCCACCTCAGCCTCCCGAGTAGCTGGGACCACAGGCACACGCTACCATGCCCAGATA
ATTTTATTATTAATTTTTGTATAGAGATGGGGTCTCCCTGTGTTGGCCAGGCTTTCTTGA
ACTCCAGGGCTCAAGTGATCCTCCCACCTCAACCTCCCAAAGTGTTGGGATTACAGGCGT
GAGCCACCATTGCCCAGCCTTAAGAGTGTGTTGATTTTCATTCAATTTCTTATATATATTAT
TTCTGTTGGGGAAAAAATTTCAAGGAAGATAAATAGTAGGCTGTTGGTACATTTCTCAAC
TTACTTTATAAAGCTTTTTAGATATATAAGGTTAATTTATGAAGAAAATCATAAGATACAC
AATTTAAGATAATTTTTTAATTTTATTTTTTATTTGTTAAATAAATTTTCTCCTTTCA
GGTGTCACGTCCTTGAAATTTTAAATGTCTCAGTGACTCTCTACCTCTGGCAGATGATGT
TGACCTTCAGCATGTAGCATCAGTAAGTACTCCTTTACTGGAGCTGATCTGAAAGCTTT
ACTTTACAATGCCCAATTGGAGGCCTTACATGGAATGCTGCTCTCGAGTGGACTCCAGGC
AAGTTATATGAGGAAGTTGTTATGACATTTTATGAGTGATAAAAGAAGTACAATGTCAAA
ATTTCCACCTTAAAAAATGCTATTTTTTAAACAACCTTTGGTAAAACGTATAGAAACATA
AATTTACCTTTAGTTGAATGTTCCATAGTTGGAATATGGGTTTTGCAGAGAATTTATAAT
TATGAAGTTTGATGTCTGTTTCTTTAACATTACCTTAATATTGGCAAAAACATGTTGGTG
TTTGCAAGGATATTATTTAAATTGGGATACCATGAATTAATACTACAAACAAAAATAAT
TAGAGTTTTTGTGTTGTTGTTGTTACTTTAACTTTTAAAAATAATCAGTTAAAGTTGTTGTT
TTGAAGCTCACATTGTTCCAATCTGGCCAATAGGAGCCCCTTTTGTATGGCTCCTGTATC
TTTATGACATGTCCTCATCATTTCTTGAATCACTTCCTCACTTCAGATACAGTAAGTTAT
TCTTGGCCAGGTGCAGTGGTTCACGCCTGTAATCCAGCACTTTGGCAGGCCAAGGCAGG
AGGATCATTGCGGCTAGTTTGAGACCAAATCATGGTTGCACAACTGTACCCACTATGG
ACAACAGAGTGGGATCTTGTCTCTGTGAAAAATTTAAAAATTAGCTGGGCATGGTGGCAC
ATACCTGTAGTCTTAGCTTCTTGGGAGAGGCTGTGGCAGGAGGATCGCTTGAGTAAATCC
AGGATGCAGTGAGCCATGCTTGTGCCACTGCACCTCCAGCATGGATGACAGAATGAGACCC
TGCCCCCAAAAAAGAAAAATATTCTTGGTTTTATCTTGTACTTTCTGTATCCAGCCCTAG
CATCAGCCTTTTCTCTAAAGACAGTATTATGATTTTAATATTTACAGTAGATATTTGAAC
TGTTACATTATAGACTTTACCATATATTTTCTAGGAAGGATTATTCTATTACTCTTCTTT
ACCACATTTGTTTGGAAATGTCTACAGAACCTACAGTTTCTAAATCAGAACTCCCTAGGT
TTTTTGCTATTTTGGCAAGCCATTGAAGTTCTTCCCTCTCCCTTTACTACCAGAAAGGTGT
GTATTTGTAGAGCTCTCTATAATGAGAAAGCACTCTATAACATGGTTGATTTCATATTTT
GGAGTAGAAAAGTATGAATGGAAAGTCAGAGACATAAAAAATAAAGCCAGAGGTCTGAGT
CTTAGCTTCATTACAGACTTTCTTGGGGGATGGTTGGTAAATTATCTACACATTCTATCT
TGTCCTTTATAATTTTAATAGTTAAATTTTTTACCATGTGCCTCAAACCGTTAGAGAATTA
ATGAGCTCTTTGAAAAATGCTTCTAAGTTTCTTGTATTGCTCTAATAGAATGCTATCTAT
GTTATTATTTATTTCTGAGACTAAAATTGTTTACATCTTTAAACTGGTTGTCTTTTGTG
TATTTTAGGATGGAAGTTCCAGCTCTGATAGTGACCTAAGTCTGTCTTCAATGGTCTTTT
TTAACCATAGCAGTGGCTCTGACGATTGAGCTGGAGATGGAGAATGTGGCTTAGATCAGT
CCCTTGTTTCTTTAGAGATGTCCGAGATCCTTCCAGATGAATCAAAATTCATATGTACC

FIGURE 14.12

31/64

GGCTCTACTTTGGAAGCTCTTATGAATCAGAACTTGGAATGGAACCTCTTCTGATTGG
 TATCTTGTGCAGTCATCATTATACAGTTCTGAAATATAAAGCTATATGTTGGTGTAAGT
 TGCAGTGATTCTCTCCTAACAGCCCCACATATTCTTCTGGTTGGTTGGTTCTTCTCAGT
 AAAATAGTCTTGTCTTCTTGCTTACACTAATTGGTAATTTGCATTCTTGTGTTAAGATTTTC
 AAGACAGGGCTGGGAGCAAGGAACCAAAGTAGCGGTGGTTGTGATTACCTTTGGTTTCT
 TTGAGGTTTCTCTTACCTAGTGGCTTTAAACATCTTTAGGAGCAGTTCCATTTTATAGT
 AAACCTAAATTTCTGTTATCATGAACAGTTGAGGATAATGAATAATTTGATACAATAATGT
 AAGAAATTCCTGAAAACAAAGTGTTATCTGTGATACTTTTGCTGCATAGTAAGCACAAATG
 AAGTGTAAGTATGTTTCAACAGGAAAGTGTTTGTGATTAAATGTGGGCAGTATCACTG
 TTCTACTAGCATTCACATCTCTTCTAAAAATTAATAGTGGTTCACTGTAATTTTATTGG
 TACATGTAAACATCTGTACATGTGTTTGGTTATCTATATGTTTCTGGTTTGTACATT
 TGCTTTATTAATTTAGGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTCACTCTATCATC
 CAGACTAGAGTGCAAGTGGCACAATTATGGCTCACTGCAGCCTTGACCTCCTGGGCTTAGG
 TGATTCTTCCACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACATGCCACCATGCCCA
 GCTAATTTTGTATGTTTGTAGAGACGAGGTTTACCATATTGCCCAGGCTGGTCTCAA
 ACTCCTGGGCTCAAGCTATCTGCGTGCCCTTGACCTCCCAAAGTGCTAGGATTACAGGTGT
 GAGCCACTATGCCTAGCCTAAGCTCAGACTTTAAAAATATAAAGCAATTCATTTTATTCT
 CCAAGAACAGTAAGGTGGTGGTTTAAATTTAGTCTTTAATTTCTGTTTAAATTTATTCTA
 TTTAGAAATGTCCAGAACTTAGTATAACTTTACTTTCTGAAAATGAAGAAACCTGTCC
 TTGGGCATTAGTGTGTTGGATTAAAGCAACAAAGTTAAAAAACCTACCCTGTGTTATGG
 CAATTTTCACTTGATGGTGGTTCTATAACACAGGTATCAGTGAACCTTTATAAAGATGA
 ACAACTTTTCACTTGCTTAAATTTTCACTTAAATTAACATGTATACTTATCTATGTTAATGT
 TTTATTGCTTAAATTTTAAATTTTATATTGTTTGGTAAACAGATAGTTTTTCTCTCCCC
 TCTTCTCTTCCATCTTCTTCACTTCAATTTTACCATGCAGAGCTCACAATGTCTCTCTGCA
CCAAGCTCCATGACTCAGGATTTGCCTGGAGTTCTTGGGAAAGACCAGTTGTTTTACAG
CCTCCAGTGTTAAGGACAGCTTCACAAGAGGGTTGCCAAGAACTTACACAAGAACAAAGA
GATCAACTGAGGGCAGATATCAGTATTATCAAAGGCAGATACCGGAGCCAAAGTGGAGTA
 TGGCTTTTTCCCCCTCATTATAATTGTTAAACTTCTTAAAAATTGTTTCAACCTTTTGA
 TATATATTTCTTTGACTTATAAACGAGCTATATTTATAAACAAGGGACCAGAACACATTA
 ACTCAGTCATGGTTATGTGCTTCTTGTCTTCAATGTTTCTTATCTTATAAGGAAGAGA
 ACGTATGGTCTCTTGAATAAAGTACAAATAAGAAGTAACAAGTGGACTACCACATTTTTT
 TTTACATCCTTAATTTAACTCTTCTGTCATTTCTTTTTTACTTAAGGAGGACGAATCCA
TGAACCAACCAGGACCAATCAAACCCAGACTGGCTATTAGTCAGTCACATTTAATGACTG
CACTTGGTCACACAAGACCATCCATTAGTGAAGATGACTGGAAGAATTTTGCTGAGCTGT
 AAGTAACAGATTCTGTTTGGGAAGTACAGCTACTATTACAAGTGACATAGTATTACACTT
 AAACCTTTAAAGTTCGTGTTTAAATATAAATATTTTGAATATTTAAAGCTAATTCAA
 AAATATGTGTCGTAGCTATGCATTAACCAACCCCAAAATGTGAGAAGTACAGAAGTCAA
 ATTGAGTTTTTCAATTAACAGTTCAATTTGATTATATTTGAATTATTCATAATGGACTCATT
 TAATTTTAGTAACCTTTGGGCTGGGTGCTGTGGCTCATGCCTGTAATCCAGCTCTTTGGG
 AGGCCAAGGCAGGTGGATCACCTGAGGTGAGGAGTTCGAGGCAAGCCTAACCAACACGGG
 GAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGGCATGTGCCTGTAG
 TCCCAGCTACTTGGGAGGCTGAGACAGGAGAATTGCTTGAACCCAGGAGGTGGAGGTTGC
 AGTGAGCCGAGATTGCACCACTGCACCTCCATCCAGCCTGGGCCACAGAGCGAGACTGTGT
 CTCAAAAAAAAAAAAAAAAAAATTTAGTAACCTCGAAGAAATAAGAAGGAAAATTTAAAGT
 TGAAAGTGATTCTAATGTATAGTTTATAAATTTTGTATAAATAACCTGTTTGCCTT
 CAAAATAATTTATATTAATATTTTATTGACCTCAAGAACATTTAAATACATTCAGATTTA
 TTCATTTGTGGACCACATTTGTTATACATTGGATTTAAAGGATCCTTGCAATTGAGTTTA
 TGGCCACCTATGCATCTGAGACCCATGGACTGGGAACCATTCTAGGTCAATGATTCAGTG
 TGATTCAATTTAAGAGATGTTTATTCCTGGTCTTTAGAAGCTGCTACCTTTTGTATCTA
 ATTTTGCAGTACTTTGAAGTATGTATGTATGTGTACATACGTTAGTGCTATGTATTTATT
 AAAGAAGAATCAGAAAACAGAGGTAAGGAAAAATAAGGAAACAAATTTCTGTAAAGCCCA
 CCACCTCCCAAAGCATATTTGTTTATATGCTTATATATGTTTCTTATTATGGTAAGAAC
 AGTCTGTACATATTGCTATATAGCAGTCCCCCTTTATCCACATACATCCTGAAAATGTTT
 TTACATTTTAAATGTTAACTACTTTATTGTTTTTAAATGTCAATTTTATAGTGTAGCTATG

FIGURE 14.13

32/64

CCACAATATCCAATTTTTAGACATTTAAATTGCTCCCAGGCAATGTGGTAATGAACATTC
 TTGCAGCTGAATATATGCACATATCTAATTGTTTCTACTAGGATAGAGGTGGAATTGTATA
 ACAGGGAGCTCACATTTTTTAAGGCTTTTGAAATGTATTGCCAAATTGCCCTGCCAGATAT
 ACTGCACCATCACTAACATTGTGTGTTCAGTATTTTTCTAACTTGGCCCTTTTGATTTT
 TAGAAAAATGATATCAATAATTTACATTTCTTTGATTAAAGTGTAGAAGTTATAATTTTT
 CATATTATTCATTGTCTATTTGTATTTTATCTTTTCTAACTTGTCTCTTCATCCCCTTTGC
 TCCGTTTTCTATTGGAGTGCACCTTTATTTGTAAGAATTCTTTTAAATTTCTGTGACTGG
 AATTTTTTTTTCTAGTTTGTATTTCCTGTTTCTTTCTAAATATAATTGTGTTTGCCA
 ACAATCCATTATCTTTTGTTTTGTAAATGGTAGTATTTATACATATTAAATTATCTCTTTC
 TTTTTTCAGATATGAAAGCTTTCAAATCCAAAGAGGAGAAAAAATCAAAGTGGAAACAAT
GTTTCGACCTGGACAGAAAGTAACCTTTAGCATAAAATATACTTCTTTTGTATTGGTTCT
GTTAAGTTTTTGTATGGCTTTTCCATATGTTGTAACAGGAAAAAATGGTGTCTATGAAT
TTCTTCTTAATTTAACAAATTTGGTTAATTTATAAAATCACAGATTGGTAAATGCTATAA
TTATGTAATGATCAGGATTGAGATTAATACTGTAGTATAAAATGGGACATTATAACAGAT
TCCATATTTTATTTCTTAAATCTAAATTCAGTCTTTAATGAAATAATATTAGCCAAATG
GTGGAATAATTTATTTCTTTTGTAGGAAAAAGATAATAAAGAAATGTAATTAAATTTAAAT
TCTTGGAATTTCCAGTTGTATATTCTATCACCTTTGTAGCATTGACAAAATTTATGCTTA
GCAGCTTCTTCACTGTTTTGAAATAAAAATATCCTATTACCTACTGATACAATTATCTGTT
 CTTTGTATATCAAAAAATGTGAAATTTACACATAATTCAAATACATTTAATTATCCGCTC
 AACCAGAAATGAAATCACATCCCTCTACTATACTACATCCAGCTCCAAGCCCAAGATATT
 TAAATGACATCCATTCTCTCTAGTTCCAGTTATGATTTTATCTTGATATTCTCTCATA
 TATGAACATAAATTATAAAGTTAGCCACCATCAATACAATCTGCGTATCTAATATCTTAAC
 TATATAGTAATGGGGTAAGGGAACAGCAAAAAGGAGAACATTAAATTAATATAACAAGTA
 AGCCTGGGCAACATAGTGAGACCCCATCTCTTAAAAAATAAATAGCCATGCATGATGGT
 ATGCCTCTAGTCCCAGCTACTTGGGAGGCTGAGGTAGGAGGATCACTTGCTCCCAGGAGG
 TTCAAGGTTCTAAACCAGCAAAGCTCAGAATCCCAGGGGATAGAAACAAAGACTTAGTGG
 ATCACTAGTATTAACTGAGACACGTCACCCTGCATTGCACCTTTGTTTCTCAGTTCTTTG
 ATGAAATCACTGAGCTGACATACCTGCCCTCTTTTACCATAAAGTGAGTTTCATGATCA
 GAAGCAATGTCTATGGGATAGCCTAACAAACAATGTAAAAACCATTTAGTAAGTTTCATGA
 AGGGTGGTGGTGGTAAAAATTTGGAGAACATACAAAACAAATACAATTCGAAGGTGTGTC
 CCTCCAGGAAGGACAAATTGTGCTGCTGCTCTGTGATAGAAGAGGATCAGATGTAATCAA
 CCTGCCGTGAGACTTGGGCTGTTCTCTCTGCTGGGTGTGGACTTGCCTGGTTGGTCACTGCT
 GCTGACAAGTAGGCTGTCAATATAGCTGGGTGTGTCATGTCAGCTGTGGTGAGGGGAAGT
 CCACATTGTGGAGGCCACATCCCTGCACTCTTGGCCAATTTGACCATGAATCTTAAGCAC
 TGGGGTGGCTGGAAAAGACAGCCGATTGACATCCATACAGAGGTCACTTGACCACTTGA
 TTAGTATAAGCACTGAAGGCTTTTAACTGAGCATTACATAGGACACAAATATTCTGATT
 CTTTGGGCCCATTCCAAGAACTCTGGGCATACTTTTCTCCAGACCTCATACCCAGTTGT
 GTTCTTTCAAATTTCTGGTCATCTGGTTATGTTATTAGCCACTATCTGTGAATCAGCAT
 AGATTTTATATCAGACATCTCTACCTCCTGACAGAATGGAGGAGATATGTTACTTAACA
 ATTCTGTTCCCTTGAAGATTTCTGTCTCCACTGTTTGTAAAGGGCTACTCCCTCAATGT
 AGCAGTAATGCTTTCACTCTGATGGGAAGTCACAGTGGAATTCTGGGTCTCCAAGAATTA
 GTGTTAGTGATACACAGTGTCTGATAATCCCCAGAGTGTCTGGTGCCCTTGGATCCTGT
 GAAGAAGGCTTGGAGAAAAGAAGATTGATGGCAAGAAGTGTGATGTGATGACAGGGCCT
 TTTCTCTGGCTCTTCATTCTTAGTCTGACCTAGGTGTGAGAATTAGGTGAGGGGCCATGA
 CTATATTGTGGTGACTCAAACCAGGCCTTTGTTTACTAACTGGGAGATTTTACATTGTA
 AGAATCAAGTAGGATCTTTGCCCATGTATTTTGGTCTTAAGAACACAAATGATATGGCTC
 CAATGACTGGAGGAACACCAGGGTCTTGGTCTCACGCTGATTTAGATAAAACGACTGTC
 AGGCCTCTGAGCCCAAGCTAAGCCATCTCCCCTGTGACCTGCACGTATACATCCAGATG
 GCCTGAAGTAACCAAGAATCAAAAAGCAGTGAAAAATGGCCTGTTCTGCTTAACTGA
 TGACATTCCACCATTGTGATTTGTTCTGCCCCATCTTAACTGAGCGATTAACTTGTGA
 AATTCCTTCTCTGGCTCAAAACCTCCCCCACTGAGCACCTTGTGACCCCCGCCCCCTGCC
 CCTAAGAGAAAAACCCCTTTGATTATAATTTTCCACTACCCACCCAAATCCTATAAAATG
 GCCCCACCCCTATCTCCCTTCGCTGACTCCTTTTTTCGGACTCAGCCCGCTGCACCCAGG
 TGAATAAAACAGCCTTGTGCTCACACAAAGCCTGTTTGGTGGACTCTCTTACACGGAC

FIGURE 14.14

33/64

AAGCTTTAGTAGAGATCTCAAAAATGGTTGGATGGTAGCAAATTACTAAGAACTCTCAAA
GTTTCTAAAGCCTTAGTTTTAGCTTGCTAGAAAACCTATGTTGAGTATTATGGCTAGTTC
CATAGTTGAGTTGGGAAATGTCTTTGAGGAGACACTTTTTCACTTTGTATTTCATCTGTAC
ATTTTCTGTTACTTGCATTCTGTCTATGCTCAGGCTATTAGAGCAGGTACATTTTTATAAC
TGGAATGTTTATGTGTAGTGAAGCTCTGAGAGGACTTTGCATTAGATCTCAGCAGCATAA
TCAGAAGGTTGTCCTTTGTCTCAGCAATTTTTAAGCTAATAGTAGCAGAAAATTGCAGTGG
AAATAGACTGCTTTGCCACAACATTTCAGAAAATCATTTATCTTTTTATTGCAGTTCCTGT
CACCAAACAATACATTTTAGTACTTCTCAAATTGCAGAACTCTCATAGGGCTGGGAAAAAT
GCCTGTAGACACATACTATGAATGTGCTAATGTTTTTTGTATTTTCATAGCCCATC
AAAGCTCCTGAGTCAGTTTCCACTATAATCACTGCAGAATCAATCTTCTACAAGGTAAGC
TTTTGTAGAGTTACTGAAGGAAGAGTTGGGCCTAGTGGGTAATGTGCCACTAAAATGTTG
GATTAGTCTAAAGGTCTCTGCTACTCTTTATTTGTATAAGGTGTGATTATACTTTTGT
CCCTTCTTAGCTGTTTTCCCCCATAAGTGGCTGTTATTAAAACATCTCATCTAGAGCTGA
AGTGGGAGGAGAAAGTGCCTACTGACACATGATGTGAGGATCTTAAGTATTTTTTTTAG
TGTAAGATTGTAGGAATTATTCTTAAATGCTGATTGTATAGTGTGGAGCCATGGAAGACT
GAGCCGTTAGTGCGATGGCATTGAAGAATGAGAAGGACAGAGACAGGATTTGGACTAGTA
GAGGTTGTCGACTGTGGTGTCAAATGGGTAGAGTAGGCCCAGAGATTCTAAAATGCCTTT
AAGTGGAGTTGAGCTGAGTAAGGGCAGTAGTGAGGATTAACACCTACTAGAAATTCATAG
TGAGAGGAATTCCAAGATGTTTTGATAAAAGAATGAGGAGGTCAGGTTTCCAGGGCCAA
AGTCCATGAACATCTGATACCTCAGTGAGAGAAGTGACAGATTGTTGTGTTTAAACCAGA
AGTCTTAGGAAAAGGAATTAGAACATAGACCCCCAAGGCTCGGCAGGCCTGGCAGGCAC
GGCAGCAACCATTGAAGGCTATTTGGTGTTCGGGATCTGAAGTGTCAATTTAGGGGACAG
TGGTGTGAGTTAGTACTTTATACTTGACCCAGGTGGACTGAGAACTCAAGTGATGATGC
CCTTAAGTATACTTTTTTTTAAAGCCACAATCTATATAGTGAAGTCTGTTCCCTCCCAAC
AGGGGTACACTGGCATTCCCTCAGCAGGGCTGGGAAAAACCAACAACAAAAAAGTCTGTA
CACAGGCAAAACATCTCTTATTTTTCCAACATTTAATACATTGTTAATAAAATATCTAA
AGTTTAGCAAACAGTTGCTGTGTATCAGTGGCTGAGCATTTCATGCTTTATTTTCATTC
AGTTCACTCTATGAGGTGGATACTACTATCCCCATTTCTAGATGAGAACATTGAGGCAC
AGCGAGGTTAATTAACCTGTCCAAGATCACATAGCCAACAAGTCATGGAGTGAGGCAGTC
TCATGCCAGAGCTTAAGCCTAGAGCATAGTTCCCTGGCTCTACAGCTTTAGCAAGTGACTG
GCTATGTGACGAGGACCAACCTCTCTAATGTCTCATCTGTAAAATAGGAATTGTAAATAG
TTACTACCTCAGTGGGTCAAATGAAATCATATGTGTTAAGCACTTAGCAGAGTAAGCACT
CAATGAATAGTAGGAGTTATCACATCTTCGTATTTGTGCATTACCTTCACAGTTTACAGA
TTAAGGCCAGAAGCAACTTGTGAGCTACGGGTTTAGTGTACTAACAGTTTCCATGTGTG
TCTCCATGGAAGGGTGTGTGGGACCTGTTATTGTGACTGTCTGTACTTTTCGTATTGTTGT
CTGCCACCCATGTTTATTAATGATAAGGACAATAATGCAACAAAGTAGTCAAGTAATGT
TGCAAATGCCCAGTATTGTAGTGGCTATCACAGCAGTGCCACTGGCAGGCAGCACCATGG
TGGCAAGTTCAAGAGGTCAGTCCAGCCACTGAGCTAGAGCCCAGATCAGGCATGCAAGA
GGAGCCTGAGTGGGAGCCACTGGGGATCACGGCCAAGAGTGTGACCACCCAAGACCCAGA
ATGGCTGAGTGGCCTCCCTGGAGCATGGCAGTGGCAGAACAACCTCCATGAACCTCAGATCT
GGTGATGCCTAAACTAGTGCTGTTCTCGTGTGGACCCCTTTTCTCTACCAGAAACCTTGA
ATCCTCTCAGCAAATGAGGAGACTACTCAGATCAGTGACTTAGTCTGTTTGGTGTATATA
TATGTGTACACAACACAGCACATATTAATAAAATACCTACTATGTGCCAGGCAGTGCCTAC
CACTGGAATCTTTCACAAAGACATTGTTTTTACTTTGCATTTCTGCCTTTACACTATGAA
AGTAGATGTTTTGGATTATATTCATTACGCATACATTTGAATATGCTGTGTTATGCATA
GTAAGCCTATGATAAGCAAGTATTCTCATTTAGAATTTGGGAATATTGATTATACATGTG
GACAAACAAACCATAAATGCAAACTATTTATATGATAAATACTTTGGACTGATGGCTGG
GAGGAAGGACCAGCTATTGATGGGTAGGAAGTAGCAAGTAGCGGACTGTGGCCTGCATAG
ACCAGACCCATCCGTAGTGATCCAGATGAAACAGCCACCCTCAGACACTTGGATAAAGGG
TCCACCAGGAAAAAATCCTGGCCTATCAGGTGCTATGTTACAGTTCAGTTACTGGAAGT
ATTTCTCAAAGTGTTTTTATGGTTGAGGTACACATTCCTACAGCTTTACCTGCTGCCA

FIGURE 15.1

34/64

AGTCCCTGTTTCAAGGGAAGCAGCAATGAATTACACTGTTCCCGTAGTCAAGGACAGTAT
 ATCTTACCAAGAACTATAACCCACTTAAGGAGGTGCTGGATGTCATAAAGATTTGGATCAA
 CCATTATGGGTGTTTCAGAGGAGAGATTATTTCCAGCTCAAGACCCAGGGAAGAGGACATA
 GGATGGATACCAGAGTTCATAGGGAGGATTTAACACAGGACATGTACACATTAGTTAGTTG
 GGTATAAAGTGGAAACAGAAATGAATGAGACACAAAGCCTTGAATGCCAGAAATACTAGTA
 GTCCTGTTGTGGAAGGATATAAACTCAACTGGGAGTGGAAAGAGAAAGGCAGCAGTGAGT
 CTAGGAGATGTACAGTAGGTTGAGGTAAACATATCCTGAAGACTATAATCCAAAGATTAT
 TTTTGGTTTGAATTTGTTTTGGTTTGAATTCATGGTATCTATTTTCTTTGAGTGGATGGT
 TGGGGAGGGTGGCATGTAGAATGCATTCTTACCAAATCAGCATGATTTTCAAGACAGTAC
 AGAGAAAAGACTGCTGAGCTGATGTAGGAGCTTTGGCTGCAGTCTCTATGGCTTTCAGCA
 AGCCGTTTAACTTACTACTGCTTCATGACTGTGGCTAACAAAGTAGGGATAGTACGGAG
 CACAGAGGATTTTATAGGGCGGTGAACTATTAATACTCTCTTTGTATGATACTATAATGG
 TGGGTACATGTCAATTATACATTTGCCCAACCCACAGAAATACACAGCACCAAGAGTGAAC
 CCTAATGTGAACTCTGGTCTTTGATGATGCTATGTACAGTGTACGTTTCATCCGTGTAACAA
 GTGTACCACTCTAGTGGTGGGAGGGGTTATTGATAATAGGGGAGGATGTGCATGTGTGGG
 GGCAGGAAGTATATGGGAAATCTCTCACTTCTGCTCAATTTTGTCTGTAACCTTAAACC
 TCTGTAAAAAATAAAGTCTATTTTTTAAAAAGTGGGGATGGTATTACGGCAATATAAAAT
 CAAAATACTTTATGAACAAATCTTTTCTCCAGATGTAACTGTCAATATGCACCCTCGT
 ATGTGTATGTATAATTTTCATTCAAACGTGAAACAACTTTAGAAATTGGCACCAACATAT
 AAACAGTATACATTAGACTATCTCGAACACCTTTTACTGACCCTTTGAAAACCTTGCTT
 ACCTATTAAGGTTTCATTCTAGCTGTGATGTTCTATTTTATTTTCAATGTGGGATTATC
TTCTGTTTCCCCAGGGAGTATATTACCAAATTGGTGATGTTGTTTCTGTGATTGATGAA
CAAGATGGAAAGCCCTACTATGCTCAAATCAGAGGTTTATCCAGGACCAGTATTGCGAG
AAGAGTGCAGCACTGACGTGGCTCATTCTCTACCTCTCTAGCCCCAGAGACCAATTTGAT
CCCGCTCTATATCATAGGTAAGTTTGACAAATGGCACAGGTTTTTTTTTAACTTAGTT
 AACTCTCCAATATTATGTAAAAGAGTGTGTTAGTCAGCTTGGGCTGTCAGGACAAAATAT
 CACAGACTGAGTGGCTTAAACAACAGAAAGTCACTTTCTCACAGTTGTGGAGGCTGAAGT
 CCAACATCAAGGTGCTGGCAACACGGATTTCTGGGGAGGCTTTTCTTCTGTCATATAGA
 TGGTCACCTTCTTGCTGTGTCTCACATGGCCTTTCATGGAGTGAGAGCTCTTTGGTGTA
 TCTTCTTATAAGGACACCATTTCTGTGAGATGAGGGCCCCACCTTATGGTTTCATTTAA
 CCTTAATTGCCTCCCTAAAGGTCTCATCTCCAAGTACCATCACATTGGGGATTAGGGCTT
 CAACATATAAATTTGGAGGGTGGCGGGGGGGGATGCAATTCAGTCCATAACAAAAAAGC
 ATGAGTATTATTAAGTACAAAAAATTAGAGAGCTTTATAGAAAAATATGAGGCATTTTAT
 GTAGCTGGAGTGTGAGTGCTATCAGTTATTTTGTAGTTAGAGCAATGTGCATCTACTAAGA
 AGTGGTATGGATAAGATTTTTTGGAGTGACCCAGGGTTAACTGTACTACAAGAAATGTA
 TTGCTCAGGAAGTAGGTTATTTAGGTTACTTATTTATACAAACCTATTCAAAAAATAATTT
 AGGAAAGAACTATCCCAGTTATCCCATCTTGCAAATCTCAATATGTGTGCCTCTGCAT
 GCTACACATGTCATCTTAGGCCTTTATAGTATAAAGGCTGATAGTTGAAATGGCAGCTGC
 TGTGCTTTTGTAAATTTCAAAGCTGCCAAAACAGTTGTGAGATAGACTCACAAGAATTTA
 CTGATTAATAACAATTTTTAAAGTTTTTCAGATTTTTTACAGTTACTTCAGACTTTTTATCTT
 TCTGCAGTGAGCATGCATCATTACTTTTGCATCCTGAGAACAAAGCATAAGTGTGTTTTTG
 GAGAGAACTCCAGGGACAAAATAATATACCACTGTTATTCTCACCTATATGTCAAGTTTTGA
 TACATTACCAACAATTCTAGCCTTCTGCTTATAAGTATATAGAATTTTTATTTACCTTA
 TCTATGGATCAGGATCTCAGCAGAGGCAGTGATGTATCAGAATCACCTTCGGGATTCCTC
 TACTGCCTCTCTTTCTAATCCCAGATTCTGATATGCATCCTTGTCTACAGCGAGGCA
 GCATGGCATGAGGTGAGAACACCAAGTCTGGAGCCAGACTGTCTAGGTTACAGCCTGCC
 ATTTACCGCCATGTGACTTTGGCAAGTTTCTTAGTCTCTCTGCTCACTTTCTCTCATA
 TGTAATGGGAATAATAATAGTGCTACCTCAGAAGGTTGATGTGAGGAATGAAGGTAT
 TGATACATGTAACTTAGAGCAGTGTTGGGTACAAAATAAACATGATGCAAGTGTTCAATC
 ACTGTTTTTGGGAGAATGCCATATTCTTTAAGCCGTTAAAGAAAGAAAAAATGATTAAGAA
 TAATTTCAAAGTAATGCATGTTTCAAGGGCTAATGCCAGGTTGCTCCCAGAGTGGTCTCT
 CCCAGTGTCTAGAAATTTTAAACATCTTATGAAAATGATATATATGGTCAAAAAATGTATTT

FIGURE 15.2

35/64

AACCTTTCCCTTGGCTGCCTTCCAGGGCCAGAGGAAGATCTTCCAAGGAAGATGGAATAC
 TTGGAATTTGTTTGTGTCATGCACCTTCTGAGTATTTCAAGTCACGGTCATCACCATTTCCC
 ACAGTTCCCACCAGACCAGAGAAGGGCTACATATGGACTCATGTTGGGCCTACTCCTGCA
 ATAACAATTAAGGAATCAGTTGCCAACCATTGTAGTTACAAATTAATACTGGGTTTCC
 AGGCCTGGTGTGGTGGCTCACGCCTGTAGCCCCAGCTATTGCACCACTGCTCTCCAAGCT
 GGGCAATGGAGTCAGATTCTCTTTCTTAAAAAACCAAAAAAACTGGATTTCAGTTCT
 CTAATATTCTTAGTACCACAAGATATGTCATAGGTATCTTTAAATGAAATTCTTAGCTGG
 AAAAGTGAATAAAAGTTTTTCTCTGCTACCTAGTAATAAACAAATCATTGTTTATTAC
 TGGTCACTTAGAAAATTAAGGGGATAGGGCCAGGCACAGTGGCTTATGCCTGTAATTGC
 AGCACTTTTAGAGGCCGAGGCAGGCGGATCACCTGAGGTCGGGAAGTGGATCGCTGAGG
 TCAGGAGTTCGAGACCAGCCTGGCCAACATGGCGAAACCCCGTCGCTACTAAAAATACAA
 AAATTAGCCAGGTGTGGTGGCATGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAG
 GAGAATCGCCTAAACCCAGGAGGTGGAGGTTGTAGTGAGCCAAGATTGCACCGCTGTGCT
 CCAGCCTGGGCAACAGAGTGAGACTCTTGTCTCGGAAAAAAAAAAAAAAAAAAGGCTG
 GGCACAGTGGCTCACGCCTTAATCCCAGCACTTTGGGAGGCTGAGGCAGATGGATCGCC
 TGAGGTTGGGAGTTCGAGACCAGCCTGGCCAGCATGGTGAAACCTGTCTCTACTAAAAA
 TACAAAAATTAGCCAGGTGTGGTGGCGCACACCTGTAGTCCCAGCTACTCGGGAGGCTGA
 GGCAGGAGAATTGGTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCAGAGATCGTGCCACT
 GCACCTCAGCCTGGGTGGACAGAGCAAGACTCCGTCTCAAAGAAACAAACAAAAAATTAA
 AAGGGATAGAATATAATGAAATATATTTTGAACCTTAAATTATATTCTATATGTGTATCTT
 CCTAGGCCAAAAGCTGTAATTTCCAGAGAGACCATTAGGAACAGGTAGTATCTATTTTCT
 CCATTATTTATTTCTAGAAACTCATAAAATGGATTGTATTTTCTATAAGAACAAAAAT
 TAATTAAGGTATAGATGACTGACCAAGGGCTTAATCAAATAAAATGACTAACAGCATCTA
 TCATAAAGCCACACAAGCCTTATGTTCTCATCTCAAAATGCTGTGACAGCTTTTTGGCT
 GCTTTAACCATAAGAAAAATGATTGGTGGATGATTTTATTAGCCCAGGCTTTTAAAAACT
 TTCATCTAGGCCACGTGCGGTGGCTCATGCCTGTAATCCCGGCACTTTGGGAGGCCTGAG
 TGGATGGATCACTTGAGGTCAGGAGTTCAGGACCAGCCTGGCCAACATGATGAAACCTG
 TCTCTACTAAATATACAAAAATTAGTTGGGTGTTATGGTGCATGCCTGTAATCCCAGCTA
 CTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCTCGGGAGGTGGAGATTGCAGTAAGCCG
 AGATCGTGCCACTGCACCTCCAGCCTGGGTGATAGAGCAAGACTGTCTCAAAAAAGAAAA
 AAAGAAAAAATTTTAATTTAATCCTTCTGTAGAAACAGGCATTGAGAACCATTCCATTGA
 TCTTAATAAAGCTGCTCTTTACTGTTTCTAGTCAAAAATGAGACTTCGATCAAACCATAA
 GATTTTATACTGCAGATAGTCAGCTTACCACAAAGCCGAGAGGAAACATGTGAGATCAG
 GCTTCTGCTTGATAGTCTCTTGACTACCATTAAACGAATATTGGGAGGTGATGAAAGT
 CATTGGTAGGCCATTAGCATTGATATCTTTAAACATCTACCCTAAACCATCTGCTATGG
 ACCCATAATAAGAGGCCTGTTGTATATGAAATTGCTCTAGAATTCAGGTGCAGGTCTTTCG
 CGGTTAAGTAAGGGAGCAACACGTAAATGGGAGAGGAGTGGGGTGTACTCACTTGCCCTC
 CTCTTTTGTCTGATTTTAACCAGCATTTTCAACCTGGGAAAAATTGCAGAATCTAAGT
 TGATTGTAATGATTTTGGAGCTGCAGCAGCTTTAACTCTTACCCTTTTCCACATAGTTAT
 GGTGTTTGAGTTGGAAAGAAACAACATATAGGTAGCTACACGTACATAATTATCTCTTAT
 TCACAAAGGGTATAGTAAAATTGATTGTAAATAACTTTCTAAGTGCCAATATTCAAAACT
 TTTGGATTAAAAATGATTTTTTACCCTGCATTTACTTTGGATGTATTTATTTTCAATTTAA
 CAATTTAAATGGGGCTCTTTAACCAAAAATGGTATTTAAACCAAAACAGTATCGTACTT
 AGAATTTGGAGTAGAGGCCGGGCACAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAG
 GCTGAGGCAGGCGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGTCAACATGAAAC
 CCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGGCGTGCGCCTATAATCCCA
 GCTAGTCTACTCGGGAGGCTGAGGCAGGAGAATCGCTGGAACCTCAGGAGGCAGAGACTGC
 AGTGAGCCGAGATCGCGCCACTGCACCTCCAGTCTGGGTGACGGCATGACTCCATCTCCAA
 AAAAAAAAAAAAAAGATTTTGGAGTAGATTATCATTAATAAGTAACAGATTTTAGGAAA
 ATCAAAAAATGGCTAATAAAATGAACACAATGTAAACATTTATTAAAAATGTAGACTTTT
 AAAATCTATAAATTGATCATCTGTTTATAAATTGGCAGATGGTTGTGTACCATCTTTTA
 AAATAAAGATTGAATTTACCCAGTGTGATGGTTCCCATTGCTTATATTCTCCTGCTGA

FIGURE 15.3

36/64

GGCCGGACCTGATATGGCCCTGGTCTGTGTTCCAGCCTTGTTTCCTCATTACCACTAAA
ATCTTTCCCCTGTATGCCCCCAATTTTTCTGGCTCTGAGTCCTTGTTCACTGTTCT
CTCCAATTCTACCTTCCAAAGGCCTTTCTTAACACCTTCGGATTCTTTCTTTGAGAACTT
TCCAGATTCCCATGCCTTTTTGGAATCAATCTCTATCCTATTGTCATCACATTTAAGTTT
CTACTTCCATCATCCTCACTCCTATCCCTTTGGTCTGGGATGACAGGGATGCTGTGTTT
TATTTACTCATCTTTGTAACCTCCACATAACCTAACCCCGGTTCTTGCTTATGGGAGATG
CTGATTGTAGGGTCTGAGTTAGATACTGTTAACTAAAATGCTTGTTGATATTTAGTTAT
TAATTCATATTAACCTTTGGCTGAACTTTTAAATTCTATTGTGAATAGTCAAGTAAAATT
TAGATTGTTACATTCTGGGTAGTATTAGATTGTTTTTAAGATTGTTTTAAACAAGATGT
TTTTAAGATGAGTTTTAAATAGTTCTCTTAACACAAATAAAGCTTAATATGAGTATTTGA
AGGAAATTATCCCAAACCATTCAGTTCTGGCTGTGAAAGGCTTTTCCAGGCCTAATAA
GTTTTCCACTTCAGCCGTAAGTAGGTGAAATCAAATGAACAATAGAGGGAAATGTATTTA
TTTGCTTTTATACACATGCATGTGTGTTGTGTCTACATATAAACATTGCACACGCTTAGAA
TGAAGTTTCTGTATGCCAGAAAAGGGAGAGGCATTTTTGTGGATTTTGTCTGGCTGCC
CTGGGGATGTTTGAAGAACTGTGCTGTTTACTTCATACCAGGTGTGTGAGCCATACCTTT
GGTAGGAGGTATACCTCCTACACCCAAGAAATATAAGCCAGGAGAAGGTCTGTGCCAAG
AGAAGGAACCCAAATGACCCACAAGAGGTGGGCCATTAATTATTGGGTGAGTGCATAAA
TGCACAGTAATTTATTTAAGCACCTCTTAATGGTGACCCACAAGGAAGATTGCTCGTAGT
AGCGGAAAGGTTTCAATAAATAAGAGAAAAAGCAGAATGTAGAAGTGTATGATAGCAA
TTCTGCAACAAGAAGCATCTTTTATAAAAGATGGAAGGAGCCAGGCACAGTAGCTCAT
GCCTGTAATCCCAGCACTTTAAGAGGCTGAGGTGGAGGATCACTTGAGCTGCAGTGACCC
ATGATTGTGCCACCACTCCAGCCTGGGTGATAGAAGTGAGACCTTCTCTCAAAAAAAAAA
AAAAAAAAAAAAAGACGGAAATTCCTCCAGAATTTTAACATGTCAACAGAGGTTTTCTGC
AGCTACTTTTTTTCAGCTTTTATACTTCGCAGTATTTTCCAAATTTTCTCTAACAAGCAGTA
TTTTCCAAATTTTTTACAATAAGCACACACACACACACGTTTGTGTCATAAGTGCCC
AACTGGTGGTGAACAACCGCTGGCTTTTAGTCTATACATATCTAGAATATTTTATAAATA
GTAGTTCTTAAACCTTGAAAGGGAGTGAATGACCAGCTGAGAAAATAAAGTCAGTGATT
TCATTATTTTCTATATTACATCATGATTCTAGGAAAGAACTTGGGAGTGACTTCCTTC
AGCTTCAGCCACTCCTGGGCCAGGCGCATGCTTAGCTCTGTGGTAAAGGTCACCAGCTTC
TTCTGCAGGGTGCCTGTATCATCTGAATTGGAGGTTTGGCGAGGGTAAGAGACTGATGTA
GGTTCAAGTTTTTCTTTCCTGTCTCCTCACTTGAAATCTGTCTTCCCTTCAGACTGCCTG
CGCTGCTGACTTAAGGCCCCAACACCAAACACAGAAGCAACAGCCTTACACAGAGTGTTT
AGCAAGCTCCAACAATTGTGTAAGGTAAAGTTTCTTTTATAGATTCTTTTCTATATCGC
TCCTAGTGGTCTGTTTCTCTGATCGAATTCTGGCTGATAACAGTTGCTGAGACTCTGAA
AGAGAAGGCAAGGAAGTACTGTTTCTCATTATAAACTGTTTAGAATTATTTGGCCATCTT
TTTGCTATGAATATGTAGTGCTTTTGATACATTTTTTAAATCAAAAAGTAATGAAAGAGAT
CACATAGGGAAAGATAGATTGGATTATTTTTAAAGTTTATATACTAAATTGAAAAGCAAA
GAATAAAATGGGAGAAACAGCTCCCTCATGTGGCTGTTGGCAGGAAGCTTCCATTCTCT
CTGTGGGCCTCCACAGGTTTGCTCACAGCAAATGGTCCGTGACAGAAAGACGCAAGGGCA
GTTGCACCCAAGATGGAAGCCACCATCTTTTCTATAACCTAATCTGAAAGAAGGGACATA
CCAGCACTTCTGCCATATGCTGTTGGGTACACAGACCAACTCTGGTACAGTGTGAACAC
AGGACCACACAAGGGCGTGAATTCCAAGGGCAGAGACCACTAGGGACCACCTCAGAGGCA
CAGAGGGACACCCTATCCAGCTGGTGGCCAATGTAAATTAACATAGCTTTTTAGAAATAGC
AATATGTATCTATAATCTTAAAGTATTAAAGTACTTCTTGATCCAGTAATTTTCAATTC
TAAGAATCCATGCTAAGAGGATTTAAATGTGGACAAAAAATGGGTATAAAAAGAAGTT
GTTAACAGTATTTAAAGTTGTGAAAAACCAGAAACAATCTAAAGGTCCAACAATAGGAAA
ATGAATTTTGATATTTTTCTAATAGAATTTTATGCTGTATCAGAAATACCATTTACAAA
TAATTTTTTAATAACGCAAAAAAAGTTTATAAAATGTTTAGTGTAACCTGGACACAAC
TACATAATGATTCTGATTTTGTAAAAAACAACAAAAACACACATATACACATGCA
TACATATGCATATAAAGAAAACCTGGAACAAACAAAATAACAAGCATAGTTGGAATTACAG
TCATTTTAATATCTTTATGCTTTTAAAAATTTTGAAGTTTGTATTACTAGCATCCACTA
CTTACGTAGTCAGGAAAAAATACAACCTTTAAATAGATATTTAGGTCCAAAGATGGTAA

FIGURE 15.4

37/64

TCTAAATGGTGTTACAGGCTGAATGTGTGCGCTGATCCCCATGCCCCAAGTTCATATGTTA
AAGCCCTGGCCCCCAAGGCAATGGTATTAGGGGAGTAGGGCCCTTTGGGAGGTAATCAGAT
TTCTACGAGGTCATGAGGGTGGAGCCCGCATAGTGGAATTAGTGTCCCTTTTAGGAAGAGG
AGAACAGACCAAAGCCTTCCTTTCTCTCCTCACTATGTAAGAAGACAGCCAGAAGGTGGC
CACAGCCAGGAAGAGAGCTCTCACCAGAACCCAAATCTGCTAGCACCTTGCTCTTGGGTT
CTCAGCATCCAGAACTGTGAGAAATGAATGTGTGTTTAAACCACTCAGGCTACGGTA
TTTTGTGTCAGCAGCCCAAGCTGACAGAGATAGAAAACAACAAGGACCCATCAGCAGAC
GAATGGATGATCAAAACGTGGTGAGGTCGTGCAGTGGGATATTATTTCAGCCGTAGAAGGA
ATGAAATTCTGATACATGCTATAATGATGAACCTTGAAAAACATGTTAATGGAAATAAGCC
AAACTTAAAAGGACAAATATTGTATAATTCCACTTATATGAGTTAGTTACCTAGAATAGG
CAAATTATGTCATAGATACAGAACATTAGAGGTTACCAGGGTTGTGGGAAGAGGGGTATT
GTGGGTACAAATTTTCGGTTTGGAGTGATTTTGAAAAAATTCTGGAAATGGGTAGTGACA
GTAGTCAACATGATGAATGTACTTAATGACACTAAATTGTACACTTAAAAATGGTTAATA
CTGGGCTGGCGCAGTGGCTCATGGCTGTAAATCCAGAACTTTGGGAGGCCAAGACAGGC
GGATCATGAGGTCAGGAGATTGAGACCATTCTGGCTAACATGGTGAAACCTGTCTCTAC
TAAAAAATAAAAAACAAATAAAAAAATAATTAGCCGGGCATGGTGGCAGGCACCTGTAGTC
CCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGTGTGACCTGGGAGTCGGAGCTTGCACT
GAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGCAACAGAGCCAGATTCCGTCTCAAAA
AAAAAAAAAAAAAGGTTGATACCTGGGTGCGGTGGCTCATGCCTGTAATTTTCAGCACTTT
GGGAGGCCAAGGCAGGCAGATCAGTTGAGGTCAAGAGTTAAGGACCAGCCTGGCCAACGT
GGCGAAACCCCATCTCTATTAAAAATACAAAAATTAGTCGAGTGTGGTGGTGGGTGCGCTG
TAGTCCCAGCTGCTGGGAGGATGAGGCCTAGGAATTGCTTGAACCCAGGAGGCAGAGGTT
GCAGTGAGTTGAGATTGCGCCACTGCACCTCCAGCCTGGGGGACAGAGCGAGACTTAGTCT
CAAAAAAAGGTTAAAATTGTAAGTTTTGTTATGCATATTTTACCATAATCTTTAAAAAA
TAGATATATAGGAGATAAAGTCAACAGAATTTAATAACCAGTTGTAAATAGAGACTGAGT
GAGGAGGATGAATTAAGGAAGACATTGAGTACAACCTTTTTGGTAGGTGAAAAACTCTTAA
AAAAATACGTGGGCAAAGATCCTACTTGATTCTTATAATTTAAAAATCTCCCAGTTAGTA
AACAAGGCTAGGTGGAGATTTGCATGTGATGTGAGGTGTGTGTTCTGTTTTGTAATGTGA
GGACTGTGAGCCATCTCCTGGACTTGAATATCCATTAGATAATTGAAAATACGGATTTGA
GAACTCAGGAGACGTGCAATGCAGTAACAAAACCTCTGCACCTAGTTGATTTCTGTCTCCT
AATTTAATGCTTTTTATGGGACAACTGTTAGGCAGGTGGGCAAGATGGACAGCCATATTT
TTGTGGGTTTTCTGGCCTGTGGGCCAGCCTCAGTGCTCACTCTGAGGTCATGTCCAACTT
AGAACACATTTCAGGCCTACCACAGTCAAGGCTCCCTTTCTCAACTCTAGTCCTCTGCACA
AATATCCGAAGCCTAGAAATAATAATCATCTGTCTTGTGTCTTGCATTATGAAAGCCTA
GGAAAGGGCCTTGGGAATTAAGAAGAAATGGAAAACTGGTCTAACTGCTGCATGCTTCAG
CTTGCAAGGGAATCACTGAAATGGGGACAGGCCATAAAAGGACAACAGAAGAGTGGCTT
CAGCAAAGGCATCGTTTTTTCAGAGCAAGCTAGAGAATCCTGCCAGCGTCCTCAGGCAGGG
CCCCTGGGCACAGAGGTTAGGCAAGGGAGTGTCCCAGCATGTTGATGCCCTGAGCATCAG
AATAATGCCATAGAGGAGCTTCAAAGAGTTCATTTAGGTTTTGTAAAGCCGAACATTTT
TAGGCAAAATAAAATTTGATTTTGTGAATAAAGCTTGTTCCTTCAACTCCAGTGCAGATT
TCATAGATTGATAGTGGCTTGTGATCCAGATAAAGAAAAACAATTTTTCAAAGATTATAT
TCTTTGTAGATGTACGGATTTAGAGACCATCTAATCTAACTCCCTCATTCTACAGATAGG
AAAAATGAGGCCTAAAGAAGTTAAGAAAATACCATGGAAATGTCACTGCTGAACTGCCAT
ACGTAGGATCCGAAGAAATTTGGGTAAATGCTACTGTGAGAAATACAGTACTAGGTCCAA
AGAATCTAATACAAATTAATAATCTAAATGTTATTTCTAAAGCATCCCTGCACATGGCTG
AACTTACATAGTTTTCATTTTCTTTCTTTCTGTTGAAGAAGAGGCAATTGGCTGGGTGCA
GTGGCTCATGCCTGTAATCCTGGCACTTTGAGAGGCCGAGGCGGGTGGATCACCTGAGGT
CAGGAGTTTGAAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAA
AATTAGCTGGCTGTGGTGGCCGCTGCCTGTAATCCCAGCTACTCCAGAGGCTGAGGCAGG
AGAATTACTTGAATCTGGGAGGTGGAGGTTGCAGTGAGCCAAGATCACGCCATTGCACTC
TAGCCTGGATGACAAGAGGGAACTCCATCTCAAAAAAAGAAAAAAGCAATCACT
AACCTGTGTTGTTTATTAAACATGACAGACTGGCATGAAGTAATTACCAAACCTGTAAACA

FIGURE 15.5

38/64

AAAAAGCTACAATCTGCCAGGCATGGTGGCTCATGCCTGTAATCCCCACCTTGGGAGGC
CAGGTTGGGGGATCACCTGAGGCCTGGAGTTCAAGACTAGCCTGGTCAACATGGTGAAAC
CTCGTCTCTACTAAAAATACAAAAATTAGCCCGCGTGGTGGCACATCCCTGTAATCCCA
GTTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGCAGTGGGGAGGTTGCAGT
GAGCCAAGATCGCACCGTTGTACTCCAGTCTGGGCCGACAGAGTGAGACTCGGTCTCAAA
AAAAAGAAAAAAGAAAAAGCTACAACCTTAATCTCAACTTCTCATAACATCATCTCTACTT
CTGATTAGAAGAGTGGAAGTGGGGAGGTTTATTACAAAAAGACTGTTATACCTTACACAC
TTCTCCCCATGAATAGTGAAGGTGTGAGTGAAAAAGACAGCAATTTTATTTTTTTTTTGA
AACAGGTTCTTGCACTGTCACCCGGGCTGGAGTGCACTGTTGTGATCACTGCTCACTGCA
GCCTCCACCTCCAGGCTCAAGTGATCCTCCTACCTCAGCCTCCTGAGTAGCTGGGACCA
CAGTTGTGCACTACCATGCCAGCTATTTTTTTTTTAAGAGATGGGGTCTCACTATATTGC
TTAGGCTAGTTCTCAAACTCCTGGCCTCAAGCAGTCCTCCGACCTTGGCCTCCCAAAGGG
TTGTGATTACAGGCATAAGCCACCACCCAGCCAGCAGTTTTAGAAATAAAGGGTGAAGG
TGCTGTTGGGGAAATATAATTTAAAAAACAAAATCTTCTCTCAACCCAGAAATCCTCTCC
ATGAAGGCAGTAGAGAAAGATAAGCTTTATTATTGAATAAAAAATTAAATGAGAATGTGAT
GCACATCACAGGCACCTTTGCTAAGAGATCACAAAGACAGAAGGAAATTTACCATTGTT
ACAGCCAAGCAGGTACAGCCCATTTACATGTATGTTTTTCGAGATAAAATAGTCCTCAACTAA
GAGAACTTGACAGCACCACTGGTCACACAGTTCATTCTAACTTTACCTGATAATTGATGT
GACCACTTGTGTTATCTAAGATATCAACTTTTCGGGGGTGGGGGAGTGTGGAAACAGGAG
TTACTTTTATAGCTTGGTGCAAGGTACTCATTAAGATTAGGCTGTTACCCTCCACAGAA
ACTGGAAGATAGGTATGCTATCTGGTAATGTTTACATTTCCAGATCCTTGAGAAAGACA
TTCTAGGTCATAAAGCTGACAAAAGGCTGATTCAGTTTTTAAATATATATATCTGTATA
TGTATTTCA

FIGURE 15.6

39/64

```

actgagagacaggactagctggatttcctaggctgactaagaatccctaagcctagctgg
|||||
actgagagacaggactagctggatttccaggccgactaagaattcctaagcctagctgg

g-aaggtgaccacatccacctttaaacacggggcttgcaacttagctcacacctgaccaa
| |||||
ggaaggtgaccacaccctcctttaaacacagagcttgtaactcagctcacacccgaccaa

tcag-----agagctcactaaaaatgctaattaggc-aaagacaggaggtaaagaaa
|||||
tcaggtagtaaagagagctcactaaaataccaattaggctaaaaacaggaggtaaagaaa

tagccaa-tcatctattgcctgagagcacagcaggagggacaatgatcgggatataaaacc
|| |||
taatcaaatcatctatcgctgagagcacagggggagggacaatgatcgggatataaaacc

caagtcttcgagccggcaacggcaacccccctttgggtcccctccctttgtatgggagctc
|| | |||
caggcatttgagccagatcaggtaaccctctttgggtcccctcacactgtatgggagctc

tgttttcatgctatttcactctattaaatcttgcaactgcac--tcttctgggtccatgtt
|||||
tggt-----ttcactctattaaatcttgcaactgcacactcttctgggtccatgtt

tcttacggcttgagctgagctttcgctcgccatccaccactgctgtttgccgccaccgca
| || |||||
tgttccggctcaagctgagcttttgctcgccgtccaccactgctgaatgccgccattgca

gaccgcgcgctgactcccatccctctggatcatgcagggtgtccgctgtgctcctgatcc
|||||
gacctgcccttgacttccacccctccggatccggcagagtgtccgctgcactcctgatcc

agcgaggcaccattgccgctcccaatcgggctaaaggcttgccattgttcctgcatggc
|||||
agcgaggcaccattgcccactcccgatcaggctaaaggcttgccattgttcctgcacagc

taagtgcctggggttcacctaattgagctgaacactagtactgggttccatgggttctct
|||||
taagtgcctggggttcacctaattcaggctgaacactgggtcgctgggttccacgggttctct

tctgtgaccacagcttctaataagagctataaacactcaccgcatggcccaagggttccatt
|| |||
tccatgactcacagcttctaataagagctataaacactcaccacatggcccaagggttccatt

cctt-gaatccataaggccaagaaccccagggtcagagaacacgaggcttgccaccatctt
| || |||||
cgttggaatccatgaggccaagaaccccagggtcagagaataaaaggcccgcc-ccatctt

gggag
|||||
gggag

```

FIGURE 16

40/64

TCCTGTGAAC CTCTAGAGGA TTTGCGCCTG CTCTTCAAAC AACAACCAGG AGGAAAGTAA	7860
CTAAAATCAT AAATCCCCAT GGCCTCCT TATCATATT TTCTCTTAC TGTTCTTTTA	7920
CCCTCTTTCA CTCTACTGC ACCCCTCA TGCCGCTGTA TGACCAGTAG CTCCCTTAC	7980
CAAGAGTTT TATGGAGAAT GCAGCTCCC GGAAATATTG ATGCCCCATC GTATAGGAGT	8040
CTTTCTAAG GAACCCCCAC CTTCACTGCC CACACCCATA TGCCCCGCA CTGCTATCAC	8100
TCTGCCACTC TTTGCATGCA TGCAAATACT CATTATTGGA CAGGAAAAT GATTAATCCT	8160
AGTTGTCCTG GAGGACTTGG AGTCACTGTC TGTTGGACTT ACTTCACCCA AACTGGTATG	8220
TCTGATGGGG GTGGAGTTCA AGATCAGGCA AGAGAAAAC ATGTAAAGA AGTAATCTCC	8280
CAACTCACC GGGTACATGG CACCTCTAGC CCCTACAAAG GACTAGATCT CTCAAACTA	8340
CATGAAACCC TCCGTACCCA TACTCGCCTG GTAAGCCTAT TTAATACCAC CCTCACTGGG	8400
CTCCATGAGG TCTCGGCCA AAACCTACT AACTGTTGGA TATGCCTCCC CCTGAACTTC	8460
AGGCCATATG TTTCAATCCC TGTACCTGAA CAATGGAACA ACTTCAGCAC AGAAATAAAC	8520
ACCACTTCCG TTTTAGTAGG ACCTCTGTT TCCAATCTGG AAATAACCA TACCTCAAAC	8580
CTCACCTGTG TAAATTTAG CAATACTACA TACACAACCA ACTCCCAATG CATCAGGTGG	8640
GTAACCTCTC CCACACAAAT AGTCTGCCTA CCCTCAGGAA TATTTTTGT CTGTGGTACC	8700
TCAGCCTATC GTTGTTTGAA TGGCTCTCA GAATCTATGT GCTTCTCTC ATTCTTAGTG	8760
CCCCCTATGA CCATCTACAC TGAACAAGAT TTATACAGTT ATGTCATATC TAAGCCCCGC	8820
AACAAAAGAG TACCCATTCT TCCTTTGTG ATAGGAGCAG GAGTGCTAGG TGCACTAGGT	8880
ACTGGCATTG GCGGTATCAC AACCTCTACT CAGTTCTACT ACAAATATC TCAAGAACTA	8940
AATGGGGACA TGGAACGGGT CGCCGACTCC CTGGTCACCT TGCAAGATCA ACTTAACTCC	9000
CTAGCAGCAG TAGTCCTTCA AAATCGAAGA GCTTTAGACT TGCTAACCGC TGAAAGAGGG	9060
GGAACCTGTT TATTTTAGG GGAAGAATGC TGTATTATG TTAATCAATC CGGAATCGTC	9120
ACTGAGAAAG TTAAAGAAAT TCGAGATCGA ATACAACGTA GAGCAGAGGA GCTTCGAAAC	9180
ACTGGACCCT GGGGCCTCCT CAGCCAATGG ATGCCCTGGA TTCTCCCCTT CTTAGGACCT	9240
CTAGCAGCTA TAATATTGCT ACTCCTCTTT GGACCCTGTA TCTTTAACT CTTGTTAAC	9300
TTTGTCTCTT CCAGAATCGA AGCTGTAAAA CTACAATGG AGCCCAAGAT GCAGTCCAAG	9360
ACTAAGATCT ACCGCAGACC CCTGGACCGG CCTGCTAGCC CACGATCTGA TGTTAATGAC	9420
ATCAAAGGCA CCCCTCCTGA GGAAATCTCA GCTGCACAAC CTCTACTACG CCCCATTCA	9480
GCAGGAAGCA GTTAGAGCGG TCTCGGCCAA CCTCCCCAAC AGCACTTAGG TTTTCTGT	9540

FIGURE 17

41/64

AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTGATTTTACCCACAAGCCCAA
LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn...PheTyrProGlnAlaGln
SerSerPheArgArgThrLysAsnArgProLeuProTrpArgArgLeuAlaThrAspPheThrHisLysProLys
AlaProSerGlyGluGlnArgThrGlyHisTyrProGlyGluAspTrpGlnLeuIleLeuProThrSerProAsn

ACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTAGATACTTTACGGGTTGGGCAGAGGCCTTCCCCTGTAGGAC
ThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHisGlyLeuGlyArgGlyLeuProLeu...Asp
ProGlnGlyPheGlnTyrLeuLeuValTrpValAspThrPheThrGlyTrpAlaGluAlaPheProCysArgThr
LeuArgAspPheSerIleTyr...SerGly...IleLeuSerArgValGlyGlnArgProSerProValGlyGln

AGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCAGATTTCGGACTTCCCCGAGGCTTACA
ArgLysGlyProArgGlyAsnLysGlyThrSerSer...AsnAsnSerGlnIleArgThrSerProArgLeuThr
GluLysAlaGlnGluValIleLysAlaLeuValHisGluIleIleProArgPheGlyLeuProArgGlyLeuGln
LysArgProLysArg.....ArgHis...PheMETLys...PheProAspSerAspPheProGluAlaTyrArg

GAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTT
Glu...Gln...ProCysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu
SerAspAsnSerProAlaPheGlnAlaThrValThrGlnGlyValSerGlnAlaLeuGlyIleArgTyrHisLeu
ValThrIleAlaLeuLeuSerArgProGln...ProArgGluTyrProArgArg...ValTyrAspIleThrTyr

ACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACACTCAAAGGACATCTAAAAAA
ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu...AsnThrGlnArgThrSerLysLys
HisCysAla...ArgProGlnSerSerGlyLysValGluLysMETAsnGluThrLeuLysGlyHisLeuLysLys
ThrAlaProGluGlyHisSerProGlnGlyArgSerArgLys...METLysHisSerLysAspIle...LysSer

GCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCCA
385 395 405 415 425 435 445
AlaAsnProGlyAsnProProHisMETAlaCysSerValAlaTyrSerLeuLysLysAsnLeuGlnLeuSerPro
GlnThrGlnGluThrHisLeuThrTrpProAlaLeuLeuProIleAlaLeuLysArgIleCysAsnPheProGln
LysProArgLysProThrSerHisGlyLeuLeuCysCysLeu...Pro...LysGluSerAlaThrPheProLys

AAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAG
LysSerArgThr...ProIleArgAsnAlaValTrpLysAlaLeuHisAsnGln...ProCysAla...ProLys
LysAlaGlyLeuSerProTyrGluMETLeuTyrGlyArgProPheIleThrAsnAspLeuValLeuAspProArg
LysGlnAspLeuAlaHisThrLysCysCysMETGluGlyProSer...ProMETThrLeuCysLeuThrGlnAsp

ACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGGAACCTAT
ThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr
GlnProThr...LeuGlnThrSerProPro...ProAsnIleAsnLysPheLeuLysHisTyrLysGluProIle
SerGlnLeuSerCysArgHisHisLeuLeuSerGlnIleSerThrSerSer...AsnIleThrArgAsnLeuSer

CCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCA
Pro...GluGluGlyLysGluLeuPheHisProCysAspMETValLeuValLysSerLeuProSerAsnSerPro
ProGluLysArgGluLysAsnTyrSerThrLeuValThrTrpTyr...SerSerProPheProLeuIleProHis
LeuArgArgGlyLysArgThrIleProProLeu...HisGlyIleSerGlnValProSerLeu...PheProIle

TCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTG
SerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyVal
Pro...IleHisProGlyLysAspProThrGlnSerPheTyrLeuProGlnLeuArgLeuLysTrpLeuGluTrp
ProArgTyrIleLeuGlyArgThrLeuProSerHisPheIleTyrProAsnCysGly...SerGlyTrpSerGly

FIGURE 18.1

42/64

GAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACGCCAAAGGAACCTGAAAATCCAGGAGACAACGCT
 GluSerTrpIleHisHisThr...ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAla
 SerLeuGlyTyrIleThrLeuGluSerAsnProGlyTyrCysGlnArgAsnLeuLysIleGlnGluThrThrLeu
 ValLeuAspThrSerHisLeuSerGlnIleLeuAspThrAlaLysGlyThr...LysSerArgArgGlnArg...

AGCTATTCTGTGAACCTCTAGAGGATTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCA
 SerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys...LeuLysSer
 AlaIleProValAsnLeu...ArgIleCysAlaCysSerSerAsnAsnAsnGlnGluGluSerAsn...AsnHis
 LeuPheLeu...ThrSerArgGlyPheAlaProAlaLeuGlnThrThrThrArgArgLysValThrLysIleIle

TAAATCCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCC
 ...IleProMETAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThrAlaPro
 LysSerProTrpProSerLeuIleIlePhePheSerLeuLeuPhePheTyrProLeuSerLeuSerLeuHisPro
 AsnProHisGlyProProLeuSerTyrPheSerLeuTyrCysSerPheThrLeuPheHisSerHisCysThrPro

CCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATT
ProProCysArgCysMETThrSerSerSerProTyrGlnGluPheLeuTrpArgMETGlnArgProGlyAsnIle
 LeuHisAlaAlaVal...ProValAlaProLeuThrLysSerPheTyrGlyGluCysSerValProGluIleLeu
 SerMETProLeuTyrAspGln...LeuProLeuProArgValSerMETGluAsnAlaAlaSerArgLysTyr...

GATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCACACCCATATGCCCGCAACTGC
AspAlaProSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMETProArgAsnCys
 METProHisArgIleGlyValPheLeuArgGluProProProSerLeuProThrProIleCysProAlaThrAla
 CysProIleVal...GluSerPhe...GlyAsnProHisLeuHisCysProHisProTyrAlaProGlnLeuLeu

TATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCT
TyrHisSerAlaThrLeuCysMETHisAlaAsnThrHisTyrTrpThrGlyLysMETIleAsnProSerCysPro
 IleThrLeuProLeuPheAlaCysMETGlnIleLeuIleIleGlyGlnGluLys...LeuIleLeuValValLeu
 SerLeuCysHisSerLeuHisAlaCysLysTyrSerLeuLeuAspArgLysAsnAsp...Ser...LeuSerTrp

GGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGAT
GlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMETSerAspGlyGlyGlyValGlnAsp
 GluAspLeuGluSerLeuSerValGlyLeuThrSerProLysLeuValCysLeuMETGlyValGluPheLysIle
 ArgThrTrpSerHisCysLeuLeuAspLeuLeuHisProAsnTrpTyrVal...TrpGlyTrpSerSerArgSer

CAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCACTCACCCGGGTACATGGCACCTCTAGCCCTACAAA
GlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLys
 ArgGlnGluLysAsnMET...LysLys...SerProAsnSerProGlyTyrMETAlaProLeuAlaProThrLys
 GlyLysArgLysThrCysLysArgSerAsnLeuProThrHisProGlyThrTrpHisLeu...ProLeuGlnArg

GGACTAGATCTCTCAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTC
GlyLeuAspLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeu
 Asp...IleSerGlnAsnTyrMETLysProSerValProIleLeuAlaTrp...AlaTyrLeuIleProProSer
 ThrArgSerLeuLysThrThr...AsnProProTyrProTyrSerProGlyLysProIle...TyrHisProHis

ACTGGGCTCCATGAGGTCTCGGCCCAAACCTACTAAGTGTGGATATGCCTCCCCCTGAAGTTCAGGCCATAT
ThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyr
 LeuGlySerMETArgSerArgProLysThrLeuLeuThrValGlyTyrAlaSerPro...ThrSerGlyHisMET
 TrpAlaPro...GlyLeuGlyProLysProTyr...LeuLeuAspMETProProProGluLeuGlnAlaIleCys

GTTTCAATCCCTGTACCTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCT
ValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyPro
 PheGlnSerLeuTyrLeuAsnAsnGlyThrThrSerAlaGlnLys...ThrProLeuProPhe.....AspLeu
 PheAsnProCysThr...ThrMETGluGlnLeuGlnHisArgAsnLysHisHisPheArgPheSerArgThrSer

FIGURE 18.2

43/64

CTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAAATTTAGCAATACTACATACACAACC
LeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThr
 LeuPheProIleTrpLys...ProIleProGlnThrSerProVal...AsnLeuAlaIleLeuHisThrGlnPro
 CysPheGlnSerGlyAsnAsnProTyrLeuLysProHisLeuCysLysIle...GlnTyrTyrIleHisAsnGln

AACTCCCAATGCATCAGGTGGGTAACCTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTGTCTGT
AsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCys
 ThrProAsnAlaSerGlyGly...LeuLeuProHisLys...SerAlaTyrProGlnGluTyrPheLeuSerVal
 LeuProMETHisGlnValGlyAsnSerSerHisThrAsnSerLeuProThrLeuArgAsnIlePheCysLeuTrp

GGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATG
GlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMETCysPheLeuSerPheLeuValProProMET
 ValProGlnProIleValVal...METAlaLeuGlnAsnLeuCysAlaSerSerHisSer...CysProLeu...
 TyrLeuSerLeuSerLeuPheGluTrpLeuPheArgIleTyrValLeuProLeuIleLeuSerAlaProTyrAsp

ACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT
ThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIleLeuPro
 ProSerThrLeuAsnLysIleTyrThrValMETSerTyrLeuSerProAlaThrLysGluTyrProPhePheLeu
 HisLeuHis...ThrArgPheIleGlnLeuCysHisIle...AlaProGlnGlnLysSerThrHisSerSerPhe

TTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACTCTACTCAGTTCTAC
PheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyr
 LeuLeu...GluGlnGluCys...ValHis...ValLeuAlaLeuAlaValSerGlnProLeuLeuSerSerThr
 CysTyrArgSerArgSerAlaArgCysThrArgTyrTrpHisTrpArgTyrHisAsnLeuTyrSerValLeuLeu

TACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGCGACTCCCTGGTCACCTTGCAAGATCAACTT
TyrLysLeuSerGlnGluLeuAsnGlyAspMETGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeu
 ThrAsnTyrLeuLysAsn...METGlyThrTrpAsnGlySerProThrProTrpSerProCysLysIleAsnLeu
 GlnThrIleSerArgThrLysTrpGlyHisGlyThrGlyArgArgLeuProGlyHisLeuAlaArgSerThr...

AACTCCCTAGCAGCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGGAACCTGT
AsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCys
 ThrPro...GlnGln...SerPheLysIleGluGluLeu...ThrCys...ProLeuLysGluGlyGluProVal
 LeuProSerSerSerProSerLysSerLysSerPheArgLeuAlaAsnArg...LysArgGlyAsnLeuPhe

TTATTTTATAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGA
LeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGluIleArg
 TyrPhe...GlyLysAsnAlaValIleMETLeuIleAsnProGluSerSerLeuArgLysLeuLysLysPheGlu
 IlePheArgGlyArgMETLeuLeuLeuCys...SerIleArgAsnArgHis...GluSer...ArgAsnSerArg

GATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGG
AspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMETProTrp
 IleGluTyrAsnValGluGlnArgSerPheGluThrLeuAspProGlyAlaSerSerAlaAsnGlyCysProGly
 SerAsnThrThr...SerArgGlyAlaSerLysHisTrpThrLeuGlyProProGlnProMETAspAlaLeuAsp

ATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTT
IleLeuProPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeu
 PheSerProSer...AspLeu...GlnLeu...TyrCysTyrSerSerLeuAspProValSerLeuThrSerLeu
 SerProLeuLeuArgThrSerSerSerTyrAsnIleAlaThrProLeuTrpThrLeuTyrLeu...ProProCys

FIGURE 18.3

44/64

GTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC
ValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMETGluProLysMETGlnSerLysThrLysIle
LeuThrLeuSerLeuProGluSerLysLeu...AsnTyrLysTrpSerProArgCysSerProArgLeuArgSer
...LeuCysLeuPheGlnAsnArgSerCysLysThrThrAsnGlyAlaGlnAspAlaValGlnAsp...AspLeu

TACCGCAGACCCCTGGACCGGCCTGCTAGCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAA
TyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProProGluGlu
ThrAlaAspProTrpThrGlyLeuLeuAlaHisAspLeuMETLeuMETThrSerLysAlaProLeuLeuArgLys
ProGlnThrProGlyProAlaCys...ProThrIle...Cys.....HisGlnArgHisProSer...GlyAsn

ATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCA
IleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer...SerGlyArgArgProThrSerPro
SerGlnLeuHisAsnLeuTyrTyrAlaProIleGlnGlnGluAlaValArgAlaValValGlyGlnProProGln
LeuSerCysThrThrSerThrThrProGlnPheSerArgLysGlnLeuGluArgSerSerAlaAsnLeuProAsn

ACAGCACTTAGGTTTTCTGTTGAGATGGGGG
ThrAlaLeuArgPheSerCys...AspGlyGly
GlnHisLeuGlyPheProValGluMETGly
SerThr...ValPheLeuLeuArgTrpGly

FIGURE 18.4

45/64

LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn...PheTyrProGlnAlaGln
ThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHisGlyLeuGlyArgGlyLeuProLeu...Asp
ArgLysGlyProArgGlyAsnLysGlyThrSerSer...AsnAsnSerGlnIleArgThrSerProArgLeuThr
Glu...Gln...ProCysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu
ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu...AsnThrGlnArgThrSerLysLys

AlaAsnProGlyAsnProProHisMETAlaCysSerValAlaTyrSerLeuLysLysAsnLeuGlnLeuSerPro
LysSerArgThr...ProIleArgAsnAlaValTrpLysAlaLeuHisAsnGln...ProCysAla...ProLys
ThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr
Pro...GluGluGlyLysGluLeuPheHisProCysAspMETValLeuValLysSerLeuProSerAsnSerPro
SerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyVal
GluSerTrpIleHisHisThr...ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAla
SerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys...LeuLysSer
...IleProMETAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThrAlaPro
ProProCysArgCysMETThrSerSerSerProTyrGlnGluPheLeuTrpArgMETGlnArgProGlyAsnIle
AspAlaProSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMETProArgAsnCys
TyrHisSerAlaThrLeuCysMETHisAlaAsnThrHisTyrTrpThrGlyLysMETIleAsnProSerCysPro
GlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMETSerAspGlyGlyGlyValGlnAsp
GlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLys
GlyLeuAspLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeu
ThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyr
ValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyPro
LeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThr
AsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCys
GlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMETCysPheLeuSerPheLeuValProProMET
ThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIleLeuPro
PheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyr
TyrLysLeuSerGlnGluLeuAsnGlyAspMETGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeu

FIGURE 19.1

46/64

AsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCys
LeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGluIleArg
AspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMETProTrp
IleLeuProPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeu
ValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMETGluProLysMETGlnSerLysThrLysIle
TyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProProGluGlu
IleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer...SerGlyArgArgProThrSerPro
ThrAlaLeuArgPheSerCys...AspGlyGly

FIGURE 19.2

47/64

SerSerPheArgArgThrLysAsnArgProLeuProTrpArgArgLeuAlaThrAspPheThrHisLysProLys
 ProGlnGlyPheGlnTyrLeuLeuValTrpValAspThrPheThrGlyTrpAlaGluAlaPheProCysArgThr
 GluLysAlaGlnGluValIleLysAlaLeuValHisGluIleIleProArgPheGlyLeuProArgGlyLeuGln
 SerAspAsnSerProAlaPheGlnAlaThrValThrGlnGlyValSerGlnAlaLeuGlyIleArgTyrHisLeu
 HisCysAla...ArgProGlnSerSerGlyLysValGluLysMETAsnGluThrLeuLysGlyHisLeuLysLys
 GlnThrGlnGluThrHisLeuThrTrpProAlaLeuLeuProIleAlaLeuLysArgIleCysAsnPheProGln
 LysAlaGlyLeuSerProTyrGluMETLeuTyrGlyArgPropheIleThrAsnAspLeuValLeuAspProArg
 GlnProThr...LeuGlnThrSerProPro...ProAsnIleAsnLysPheLeuLysHisTyrLysGluProIle
 ProGluLysArgGluLysAsnTyrSerThrLeuValThrTrpTyr...SerSerPropheProLeuIleProHis
 Pro...IleHisProGlyLysAspProThrGlnSerPheTyrLeuProGlnLeuArgLeuLysTrpLeuGluTrp
 SerLeuGlyTyrIleThrLeuGluSerAsnProGlyTyrCysGlnArgAsnLeuLysIleGlnGluThrThrLeu
 AlaIleProValAsnLeu...ArgIleCysAlaCysSerSerAsnAsnAsnGlnGluGluSerAsn...AsnHis
 LysSerProTrpProSerLeuIleIlePhePheSerLeuLeuPhePheTyrProLeuSerLeuSerLeuHisPro
 LeuHisAlaAlaVal...ProValAlaProLeuThrLysSerPheTyrGlyGluCysSerValProGluIleLeu
 METProHisArgIleGlyValPheLeuArgGluProProProSerLeuProThrProIleCysProAlaThrAla
 IleThrLeuProLeuPheAlaCysMETGlnIleLeuIleIleGlyGlnGluLys...LeuIleLeuValValLeu
 GluAspLeuGluSerLeuSerValGlyLeuThrSerProLysLeuValCysLeuMETGlyValGluPheLysIle
 ArgGlnGluLysAsnMET...LysLys...SerProAsnSerProGlyTyrMETAlaProLeuAlaProThrLys
 Asp...IleSerGlnAsnTyrMETLysProSerValProIleLeuAlaTrp...AlaTyrLeuIleProProSer
 LeuGlySerMETArgSerArgProLysThrLeuLeuThrValGlyTyrAlaSerPro...ThrSerGlyHisMET
 PheGlnSerLeuTyrLeuAsnAsnGlyThrThrSerAlaGlnLys...ThrProLeuProphe.....AspLeu
 LeuPheProIleTrpLys...ProIleProGlnThrSerProVal...AsnLeuAlaIleLeuHisThrGlnPro
 ThrProAsnAlaSerGlyGly...LeuLeuProHisLys...SerAlaTyrProGlnGluTyrPheLeuSerVal
 ValProGlnProIleValVal...METAlaLeuGlnAsnLeuCysAlaSerSerHisSer...CysProLeu...
 ProSerThrLeuAsnLysIleTyrThrValMETSerTyrLeuSerProAlaThrLysGluTyrProphePheLeu
 LeuLeu...GluGlnGluCys...ValHis...ValLeuAlaLeuAlaValSerGlnProLeuLeuSerSerThr
 ThrAsnTyrLeuLysAsn...METGlyThrTrpAsnGlySerProThrProTrpSerProCysLysIleAsnLeu
 ThrPro...GlnGln...SerPheLysIleGluGluLeu...ThrCys...ProLeuLysGluGlyGluProVal

FIGURE 20.1

48/64

TyrPhe...GlyLysAsnAlaValIleMETLeuIleAsnProGluSerSerLeuArgLysLeuLysLysPheGlu
IleGluTyrAsnValGluGlnArgSerPheGluThrLeuAspProGlyAlaSerSerAlaAsnGlyCysProGly
PheSerProSer...AspLeu...GlnLeu...TyrCysTyrSerSerLeuAspProValSerLeuThrSerLeu
LeuThrLeuSerLeuProGluSerLysLeu...AsnTyrLysTrpSerProArgCysSerProArgLeuArgSer
ThrAlaAspProTrpThrGlyLeuLeuAlaHisAspLeuMETLeuMETThrSerLysAlaProLeuLeuArgLys
SerGlnLeuHisAsnLeuTyrTyrAlaProIleGlnGlnGluAlaValArgAlaValValGlyGlnProProGln
GlnHisLeuGlyPheProValGluMETGly

FIGURE 20.2

49/64

AlaProSerGlyGluGlnArgThrGlyHisTyrProGlyGluAspTrpGlnLeuIleLeuProThrSerProAsn
LeuArgAspPheSerIleTyr...SerGly...IleLeuSerArgValGlyGlnArgProSerProValGlyGln
LysArgProLysArg.....ArgHis...PheMETLys...PheProAspSerAspPheProGluAlaTyrArg
ValThrIleAlaLeuLeuSerArgProGln...ProArgGluTyrProArgArg...ValTyrAspIleThrTyr
ThrAlaProGluGlyHisSerProGlnGlyArgSerArgLys...METLysHisSerLysAspIle...LysSer
LysProArgLysProThrSerHisGlyLeuLeuCysCysLeu...Pro...LysGluSerAlaThrPheProLys
LysGlnAspLeuAlaHisThrLysCysCysMETGluGlyProSer...ProMETThrLeuCysLeuThrGlnAsp
SerGlnLeuSerCysArgHisHisLeuLeuSerGlnIleSerThrSerSer...AsnIleThrArgAsnLeuSer
LeuArgArgGlyLysArgThrIleProProLeu...HisGlyIleSerGlnValProSerLeu...PheProIle
ProArgTyrIleLeuGlyArgThrLeuProSerHisPheIleTyrProAsnCysGly...SerGlyTrpSerGly
ValLeuAspThrSerHisLeuSerGlnIleLeuAspThrAlaLysGlyThr...LysSerArgArgGlnArg...
LeuPheLeu...ThrSerArgGlyPheAlaProAlaLeuGlnThrThrThrArgArgLysValThrLysIleIle
AsnProHisGlyProProLeuSerTyrPheSerLeuTyrCysSerPheThrLeuPheHisSerHisCysThrPro
SerMETProLeuTyrAspGln...LeuProLeuProArgValSerMETGluAsnAlaAlaSerArgLysTyr...
CysProIleVal...GluSerPhe...GlyAsnProHisLeuHisCysProHisProTyrAlaProGlnLeuLeu
SerLeuCysHisSerLeuHisAlaCysLysTyrSerLeuLeuAspArgLysAsnAsp...Ser...LeuSerTrp
ArgThrTrpSerHisCysLeuLeuAspLeuLeuHisProAsnTrpTyrVal...TrpGlyTrpSerSerArgSer
GlyLysArgLysThrCysLysArgSerAsnLeuProThrHisProGlyThrTrpHisLeu...ProLeuGlnArg
ThrArgSerLeuLysThrThr...AsnProProTyrProTyrSerProGlyLysProIle...TyrHisProHis
TrpAlaPro...GlyLeuGlyProLysProTyr...LeuLeuAspMETProProProGluLeuGlnAlaIleCys
PheAsnProCysThr...ThrMETGluGlnLeuGlnHisArgAsnLysHisHisPheArgPheSerArgThrSer
CysPheGlnSerGlyAsnAsnProTyrLeuLysProHisLeuCysLysIle...GlnTyrTyrIleHisAsnGln
LeuProMETHisGlnValGlyAsnSerSerHisThrAsnSerLeuProThrLeuArgAsnIlePheCysLeuTrp
TyrLeuSerLeuSerLeuPheGluTrpLeuPheArgIleTyrValLeuProLeuIleLeuSerAlaProTyrAsp
HisLeuHis...ThrArgPheIleGlnLeuCysHisIle...AlaProGlnGlnLysSerThrHisSerSerPhe
CysTyrArgSerArgSerAlaArgCysThrArgTyrTrpHisTrpArgTyrHisAsnLeuTyrSerValLeuLeu
GlnThrIleSerArgThrLysTrpGlyHisGlyThrGlyArgArgLeuProGlyHisLeuAlaArgSerThr...
LeuProSerSerSerSerProSerLysSerLysSerPheArgLeuAlaAsnArg...LysArgGlyAsnLeuPhe

FIGURE 21.1

50/64

IlePheArgGlyArgMETLeuLeuLeuCys...SerIleArgAsnArgHis...GluSer...ArgAsnSerArg
SerAsnThrThr...SerArgGlyAlaSerLysHisTrpThrLeuGlyProProGlnProMETAspAlaLeuAsp
SerProLeuLeuArgThrSerSerSerTyrAsnIleAlaThrProLeuTrpThrLeuTyrLeu...ProProCys
...LeuCysLeuPheGlnAsnArgSerCysLysThrThrAsnGlyAlaGlnAspAlaValGlnAsp...AspLeu
ProGlnThrProGlyProAlaCys...ProThrIle...Cys.....HisGlnArgHisProSer...GlyAsn
LeuSerCysThrThrSerThrThrProGlnPheSerArgLysGlnLeuGluArgSerSerAlaAsnLeuProAsn
SerThr...ValPheLeuLeuArgTrpGly

FIGURE 21.2

51/64

TTGGTCTTAAGAACACAAATGATATGGCTCCAATGACTGGAGGAACACCAGGGTCCTTGG
TCTCACGCTGATTTAGATAAAACGACTGTCAGGCCTCTGAGCCCAAGCTAAGCCATCCTC
CCCTGTGACCTGCACGTATACATCCAGATGGCCTGAAGTAACCAAAGAATCACAAAAGCA
GTGAAAATGGCCTGTTCTCTGCCTTAACTGATGACATTCCACCATTGTGATTTGTTCCCTGC
CCCATCTTAACTGAGCGATTAACTTGTGAAAATTCCTTCTCCTGGCTCAAAAACCTCCCCC
ACTGAGCACCTTGTGACCCCCGCCCCCTGCCCCCTAAGAGAAAAACCCCTTTGATTATAATT
TTCCACTACCCACCCAAATCCTATAAAATGGCCCCACCCCTATCTCCCTTCGCTGACTCC
TTTTTCGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAA
GCCTGTTTGGTGGACTCTCTTCACACGGACGCTCATGACATTTGGTGCCAAAACCTGGGA
TAGGAGGACTCCTTCAGGAGACCAGTCCCCTGTCCTTGCCCTCACTCTGTGAGGACATCC
ACCTACAACCTTGGGTCTCAGACCAACCAGCCCAAGGAACAGCTCACCAATTTCAAATC
AGGTAAGCAGTCTTTTCACTCTCTTCTCCAGCCTCTCTTGCTACCCTTCAAACCTCCCTCT
CTCACTACCCTTCAATCTCCCTGTCCTTCCAATTCCAGTTCTTTTTCATCTCTAGTAGAG
ACAAAGGAGACACATTTTATCCATGGACCCAAAACCTCCAGCACCAGTCACGGACTTGGGA
AGACAGTCTTCCCTTGGTGTTTAATCACTGCGGGGACGCCTGCCTGATTATTCACCCACA
CTCCATTGGTGTCTGATCACGGTGGGGACACCTGCCTTGGTCACTCACCCACATTCCCTT
GGTGGTACGTCAACTGCAAAAGCAGGGGACGCCTGCTTTGGCTGCTCACCCACCCCTTC
TCTGTGTCTCTACCTTTCTCTTTAAACTTACCTCCTTCACTATGGGCAAACCTTCTGCCCT
CCATTCCCCCTTCTTCTCCCTTAGCCTGTGTTCTTAAAAACCTAAAACCTCTTCAACTCA
CACCTGACCTAAAACCTAAATGCCTTATTTTCTTCTGCAACACTGCGTGGCTGCAGTACA
AACTTGATAATAGCTTTTAAATGGCCAGAATATGGCACTTTCAATTTCTCCATCCTACAAG
ATCTAGATAATTTTTGTGGAATAATGGAAAAATGGTCTGAGATGCCTGACGTCCAGGCAT
TCTTTTACACATTGGTCCCTCCCTAGTCTCTGCTCCCAATGCGACTCATCCCAAATCTTT
CTTCTTTCTCTCCTGTCTGTTCTTCACTCTCCACCCCAAGCTCTGAGTCCTTTGAATCC
TCCTTTGCTACAGACCCATCTGAACCTCTCCCTCCTCCCCAGGCTGCTCCTCACCAGGCC
GAGCCAGGTCCCAATTCTTCTCAGCCTCTGCTCCCCCACCCTATAATCCTTTTATCACC
TCCTCTCCTCACACTCAGTCCGGCTTACAGTTTCGTTCTGTGACTAGCCCTCCCCATCT
GCCCAACAATTTCTCTTAAAGAGGTGGCTGGAGCTAAAGGCATAGTCAAGGTTAATGCT
CCTTTTTCTTTATCTGACCTCTCCCAAATCAGTTAGCGTTTACGCTCTTTTTTCATCAAAT
ATAAAAACCCAGCCAGTTTATGGCCCATCTGGCAACAACCCTTACAGGCTTTACAGCCCT
AGACCCTGAAGGGTCAGAAGGCCGTCTTATTCTCAATATGCATTTATTACCCAATCCGC
TCCCAACATTAAATAAAGCTCCAAAAATTAAATCTGGCCCTCAAACCCCAACACAGGAC
TTAATTAACCTCACTTCAAGGTGTACAAGAATAGAGTAGAGGCAGCCAAGTAGCAACGTA
TTTGAGTTGCAATTCCTTGCCTCAACTCTGAGAGAAACCCAGCCACATCTCCAGCAAAC
AAGAACTTCAAAACACCTGAACTGCAGCAGCCAGGCGTTCTCCAGGACCACCTCCCCCA
GGATCTTGCTTCAAGTGCCGAAATCTGACCATTGGGCCAAGGAATGCCTGCAGCCCAGG
ATTCTCTCTAAGCCACGTCCCATTTGTGCAGGACCCCACTGGAAATCGGACTGTCCAAC
CACCCGGCAGCCAATCCCAGAGCCCCCTGGAACCTTGCCCCAAGGCTCTCTGACTGACTCC
TTCCCAGATCTTCTCGGCTTAGCAGCTGAAGACTGACACTGCCCGATCACTTCAGAAGTC
CCCTGGACCATCACGGATACTGAGCTTCAGGTAACCTCTCACAGTGGAGGCTAAGTCCATC
CCCTGTTTAAATCGATACAGGGGCTACCCACTCCACATCACCTTCTTTTCAAGGGCCTGTT
TCCCTTTCCCCCATAACTGTTGTGGGTATTGACGGCCAAGCTTCAAACCCCTTAAAACT
CCCCCACTCTGGTGCCAACTTGGACAACATTCTTTTATGCACTCTTTTTTCAAGTTATCCTC
ACCTGCCCAGTTCCCTTATTAGGCCGAGACATTTTAACCAAATTATCTGCTTCCCCGACT
ATTCTGGGCTACAGCCACATCTCCTTGCCGCCCTTCTTCCAACCCAAAGCCTCCTTCA
TATCTTCTCTCATATCCCCCACCTTAACCCACAAGTATGGGACACCTCTACTCCCTCC
CTGGCAACCGATCACACGCCCATTAATCTTAAACCTAATCACCTTACCCTGCT
CAATGCCAGTATCCCATACCAACAGGCTTTAAAGGGATTGAAGCCTGTTATCACTTGC

FIGURE 22.1

52/64

CTGCTACAGCACGGGCTTCTAAAACCTATAAACTCTCCATACAATTCCCCCATTTTACCT
GTCTAAAAACCAGATAAGTCTTACAGGTTAGTTTACAGAACTCTGCACCTTATCAACCAAATT
GTTTTGCCTATCCACCCTGTAGCACCCAACCTCGTACACTCTTTTGTCTCAATGCCTTCC
CCCACAACCTCACTATTCCGTTCTTGATCTTAAAGATGCTTTTTTCACTATTCCCCTGCAC
CCCTCATCCCAGCCTCTCTTTGCTTTTACCTGGACTGACCCTGACACCCATCAGTCCCAG
CAGCTTACCTGGGCTGTACTGCCGCAAGGCTTCAGGGACAGCCCTCATTACTTCAGCCAA
GCTCTTTCTCATGATTACTTTCTTCCACCTCTCTGCTTCTCACCTTATTCAATATATT
GATGACCTTCTACTTTGTAGCCCCCTCCTTTAAATCTTCTCAACAAGACACCCTCCTGCTC
CTTCAACATTTGTTCTCCAAAGGATATCGGGTATCCCCCTCCAAAGCTCAAATTTCTTCT
CCATCTGTTACATACCTCGGCATAATTCTTCATGAAAACACATGTGCTCTCCCTGCCAAT
TGCGTCTCCAACCTGATCTCTCAAATCCCAACCTCTTCTACAAAACAACAACCTCCTTTCCC
TCCTAGGCATGGTTGGATACTTTTGCCTTTGGATACCTGGTTTTGCCATCCTAACAAAAT
CATTATATAAACTCACAAAAGGAAACCTAGCTGACCCCATAGATTCTAAATCCTTTCCCC
ACTCCTCTTTCCATTCTTGAAGACAGCTTTAGAGACTGCTCCCACTAGCTCTCCCTG
TCTCATCCCAACCCTTTTCATTACACACAGCCGAAGTGCAGGGCTGTGCAGTCGGAATTC
TTACACAAGGACCGGGACCATGCCCTGTAGCCTTTTTTGTCCAAACAACCTTGACCTTACTG
TTTTAGGCTCGCCATCATGTCTCCATGCGGTAGCTTCCGCTGCCCTAATACTTTTAGAGG
CCCTCAAAATCACAACTATGCTCAACTCACTCTCTACAGCTCTCACAACTTCCAAAATC
TATTTTCTTTCTCACACCTGACGCATATACTTTCTGCTCCCCGGCTCCTTCAGCTGTATT
CACTCTTTGTTGAGTCTCCCACAATTACCATTCTTCTGCCCCAGACTTCAATCTGGCCT
CCCACATTATTCTGGATACCACACCTGACCCTGATGATTGTATGTCTCTGATCTACCTGA
CATTACCCCCATTTCCCCATATTTCTTCTTTTCTGTTCTCATGTTGATCACATTTGGT
TTACTGACGGCAGTTCCACCAGGCCTGATCGCCACTCACCAGCAAAGGCAGGCTATGCTA
TAGAATCTTCCACATCCATCATTGAGGCTACTGCTCTGCCCCCTCCACTACCTCTCAGC
AAGCCGAACCTGATTGCCTTAACTCGGGCCTTCACTCTTGCAAAGGGACTACACGTCAATA
TTTATACTGACTCTAAATATGCCTTCCATATCTTGCAACCACCATGCTGTTATATGGGCTG
AAAGAGGTTTCTCTACTACGCAAGGGTCTCCATCATTAATGCCTCTTTAATAAAAACTC
TTCTCAAGGCTGCTTTACTTCCAAAGGAAGCTGGAGTCACACACTGCAAGGGCCACAAA
AGGCGTCAGATCCCATTACTCTAGGAAATGCTTATGCTGATAAGGTAGCTAAAGAAGCAC
CTAGCGTTCCAACCTCTGTCCCTCATGGCCAGTTTTTCTCCTTCCCATCAGTCATTCCCA
CCTACTCCCCCATTTGAAACTTCCGCCTATCAATCTCTTCTCACACAAGGCAAATGGTTCT
TAGACCAAGGAAAATATCTCCTTCCAGCCTCACAGGCCCATTTCTATTCTGTATCATTTT
ATAACCTCTTCCATGTAGGTTACAAGCCACTAGTCCACCTCTTAGAACCTCTCATTTCT
TCCATCGTGGAACATATCCTCAAGGAAATCACTTCTCAGTGTTCCATCTGCTATTCTAC
TACCCCTCAGGGATTGTTGAGGCCCCCTCCCTCCCTACACATCAAGCTCGGGGATTTGC
CCCTGCCCAGGACTGGCAAATTGACTTTACTCACATGCCCTGAGTCAGGAACTAAAATA
CCTCTTGGTCTGGGTAGACACTGTCACTGGATGGGTAGAGGCCTTTCCCACAGGGTCTGA
GAAGGCCACTGCAGTCATTTCTTCCCTTCTGTGACACATAATTCCTTGGGTTGGCCTTCC
CACCTCTATACAGTCCAATAACGGAGCAGCCTTTATTAGTCAAATCACCTGAGCAGTTTT
TCAGGCTCTTGGTATTCACTGGAACCTTCGTACCCCTTACTGTCTCAATCTTCAGGAAA
GGTAGAATGGACTAATGGTCTTTTAAAAACACACCCCAACCAACTCAGCCTCCAACCTAA
AAAGGAGGATAGAGCCCCAAAACCTCGCAACCAAGCTAGTAATTATGCTGAACCCCTTGG
GCACTCTCTAATTGGATGTCTTAGGTCCTCCCAAATCTTAGTCCTTTAATATCTGTTTTT
CTCCTTCTCTTATTTCGGACCTTGTGTCTTCCGTTTAGTTTTTCAATTCATACAAAACCGC
ATCCAGGCCATACCAATCGTTCTATACAATAAATGCTCCTTCTAACAACCCCAACAATAT
CGCCCCCTTACCACAAAATCTTCTTCACTTAATCTCTCCCACTCTAGGTTCCCATGCCG
CCCATAATCCCTCTCGAAGCAGCCCTGAGAAACATAGCCATTATCTCTCCATACCACCC
CCAAAATTTTGTCTGCCCCAACACTTCAACACTATTTTACATTATTTTCTTATTAATAT

FIGURE 22.2

53/64

AAGAAGACAGCAATGTCAGGCCTCTGAGCCCAAGCCATCATATCCCCTGTGACCTGCACA
TATACATCCAGATGGCCTGAAGTAACTGAAGAATCACAAAAGAAGTGAAAATGGCCTGTT
CCTGCCTTAACCGATGACATTCCACCACTGTGATTTGTTCCCTGCCCCACCTTAACTGAGC
AATTAACCTTGGGAAATTCCTTCTCCTGGCTCAAAACCTCCCCCACTGAGCACCTTGTGA
CCCCTGCCCCTCCACTACCCACCCAAATCCTATAAAATGGCCCCACCCCATCTCCCTTAG
CTGACTCCTTTTTTGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTGCT
CACACAAAGCCTGTTTGGTGGACTCTCTTCACAGGGACGGGGGTGACAACAACACGGACA
CACATGGAGTGGTTTTAAGGAGCAGAGAGTTTAATACGCAAAAAAGAAGGAAGAGGCTCC
CCTGTACAGACACAGAGGGAGGGGGCTCCAAGCCGAGAGAAGGAAACCCCATGTGCAGTG
GAAAAGTGGTTGATTATACTGGGAGGCTGGAGGAGGCGGTGTCTGATTTGCACAGGGCCC
AGGGGATTGGGTTGACCAGGTGTATCATTATGTACCCCGCAAAAAACCTGGCCCTCCCA
CCTCAGCCCTTTAATATGCAAATGTGGGTTGCCATGATGTTCTGAAAACACATGAATTAT
CTGGAGGGGGCCATGACACTTGGTACATGTGCTGACAAGAAGAGGGTGGGAATCGCCATG
GTGGCCATGTTGGGTGGACCTAGTTTTTAATAGCCTGCATTTGCATATCAAAGTTTGCTG
GCCTGGCTCTTTAAGCTGTCTTTTCTGTTAGAAAAGGAATGGTTTGGAATGGGTGAGGGT
TGCTTCTTATTACAAGAAAATTTCCAAAAACCTTTACTCTTTCTAGCTGCCAAAAAACTA
TTTCTTAATAACTTATGTATTACCATAATTAGGCAGCACCAAAGATCCCTGCAGGTCAGA
CCACTGCAATTAACATGCTGGCTTTACTGCTGATTATGGTAGCTGCATCCACCTAGCCTC
TCATATTGCAACTGCCTGACCTCTGCCACCCACGAGCCACTTATCCCCACTTATAATCA
GCCCATTTGATTGTAACATCTGCCACTTATTTCCCGACGTTGTGGTATATCCTATAGATG
AATTCATTCAACATCCATTCCAACACCACCTCTCTTGCCCTTCTATACTCTCTGGAGAGT
GAATTACTGAGTCACATGATCTTCACTGCAGTCATTTGTGGCTATGTGACATAGTTCTGG
ACAGTGAACATAGACAGAAGTCCCTGGGGCGGGCTTCCTTTCTGGGATGAGGGCAAAACG

FIGURE 22.3

54/64

GATCTCTTGATCCCAGGAGGTCAAGGCTGCAATGAGCTAAGATCAAGCCACTGCATTCCA
 GCCTGAGTGATAGTGGGAGACCTTGTCTTTAAACACACACACACACACACACACACG
 AGGGCCTTTGACCACTCTTGAGTAGAAGACTCGAGAAGAACAAGTAGAAGGCCAGAGAA
 GAACAAAGTTACTTGAAAGATCTCTTATTAAAGAGAATGTACAAGCTATGAAAAAAAAA
 AACACACACACACACACAAACCTCATCTGGAATGAAAAAACATAATGCATTTGGTTTCT
 GGTTCTTAGGCTGTTATGGAACAACCAAGAACATTATTTTGGTTTCTGAGGTGAGAAC
 TATTTTATTCCCTCAAGCACACTATGCTTATGGTTTGAGGGAGAATGAGAAATAGGAAA
 CTAGGAACAGGCTGAAATGGTCTAATCTTGACCATCTAATTCTGCAGTGTCTTATTCTCA
 TTCTAAAAGAGAATGGTTATATTGCTGTTCTAGCATAAAAAGTAATGATAAAAATAAAA
 GATCCCGTATTACCAGACAATAATCCCTAGACTGTTTAAATGCTTGGTTGAGTATTTGC
 TTATGATCTCAGACTTTAAAGATGGTCTCCCCCTATGGTGAAGCTTGTTAATTATGTAG
 GCATCATTAATGTCTGTTTACTTATCAAAATTTTATCATTGTTAGTTGTATTACTACTTG
 ACAGTCCAATTTATTTAATTGAAAAGATTGGTTAACATTTTATAGTCAAAGTAATTGTTT
 CCTGTGTTTTTTTCTGTTTAGGTTATTGGAGTGATGAGTAAAGAATACATACCAAAGGGC
ACACGTTTTGGACCCCTAATAGGTGAAATCTACACCAATGACACAGTTCCTAAGAACGCC
AACAGGAAATATTTTTGGAGGGTAAGTAAGGGGAAATTTCTTCAGACCCATTAAATGTTAG
 GAAAAATGGAGCTAAAAGAGCTGGGTGGCTCACCTTTCTCATCTGTGCTGAGAAATGC
 TGGGGCTCACCCATAAGTATCCAGCATCCCCATGGACACAGGGAATTCTGAACAAATGTG
 ATGAAACCGATGAAATGTCTGGCCTGTAGGTGGTTAGTGATGGAGATACGGGCTATATGT
 GAATCTTGATTTTGCATTTCAATTAGAGCTTTGTAATGAAAGGAAACAGTTTGTGCTTG
 CTTTAAGGATAGGTTCAATTTGCATTTCTCCGCAAGGAAGTAGTAATGAGTTACCAAGCCT
 TAGATTTTACCCCTTTTTTGATTTCTTGCTGACTTAACCTTAAATTGAATGGAAGAGTTATC
 ACAAATGAATTATCTTTTTGGTTTTTTTTTTTTTTTGGAGATGGAGTCTCACTCTGTCACCAG
 GCTGGAGTGCAATGGCATGATCTCGGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAA
 TTGTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTAAGGTGCGCGCCACCATGCCAGTTA
 ATTTTTGTATTTTTAGTAGAGACGGGGTTCCACTATGTTGGCCATGATGGTCTCGATCTC
 TGGACCTCGTGATCCGCCCACCTTGGCCTCCCAAAGTGCTGGAATTACAGGCAAGAGCCA
 CCGCGCCAGCCAGGAATGACAAATGAATTACCTTATAAGTAAATGCCATTAAGGAAGGA
 TAGCTGGAAGATGGGTTGAGGGGAATGGAGGACCACAGAACTAGTCTTATTTAAATACAT
 GTGCATGGTAAAATGATTCCATTTGACAATAGGTTAATTATCTCATAGCATAAGGAAAAT
 GCTTAACAGTCATATGCAAGATGATAAGCTTTCTTATAGCATCCAACCAAAAGATCTAGC
 CAGTACAATTTCTTTGCTATATTAGGGTTAGAAAGGCCCCAGAGGTGAACCAATTAGA
 TGGAATCCTTGAATAAAACACTGGATTAGCAGTGAACAGAAAAAGTCAGATTGCTTTCC
 TTCTTCCCATAGATGTCTCAGGGATATTTAGTTTCTCAGAAAGATAAAGAATTTAGTAAG
 CGTTTTTTTGTGCATACTTACATGAAATGTACATTATTTGAATTTCTTTAAAAAGAAACAG
 CTGCATGATAACAAAAATTGTGTTATGCTTGCTTTAGCTGGTATTTTTGCCTAGAACGAT
 TATATCGTTTCGGACAAGAAGCTATTCCCTAAGAAACAATATTTTTAATCCAGGAAGTTTTT
 CATTTTTTAGAAATTTTACTTATTTCCTAAGCAAAAGAGGGTAGTTACAGATTCACCTA
 AGAATCATGTGCTCACAAATTTTTATTTAATAATTATTCCTCCTTAAAATATATTAATCAC
 CTGACTTACAATGGTGGAACCATGAGTGCAATTTTGCCTTTATTGTCAATAACGTCTTCT
 CAGAAGTGAGCCACAAAGGTGCATAGTTCTTGGAGTTAAAGGTCTGAATTAAGACAATCC
 AGCATAAGTCTCATTAATGTGTGATTATTTTGGAGAAAGGCAAGAAGTACCTAAGAATCT
 CCCCCTCACTGTCCAGTTCCCTGTTTCATTTAAAGATTCACTGTAAGTAACTGAAAGGCT
 TTCCTTGGGAGGATTTATTTGAATCAGTCTTTCACATGCAAAGGATATTGTAGAACATCT
 CGTTTTTGTCTGGCAGGAATATGAACATCTGTTGTGAGGAAAGAAAAAGTTTCATGCAAAAT
 TACACTGCCAAAGAAGGGATGTTCAAGTTGAGAAACCAGTGACATTTCTTGTAAGTGTAC
 TATGAATCAGCGCATTTTAAATCTTCTAGATAATATATGGAAGTGCAGGAAGGTGGTAGGA
 AACGGTGTTCATTTTACATATGCGTTATTTTATTCTGTGTGAGTGACTTCATGGCACCAG
 CATTGCTGTTTTTAAATGAGGATACAGTAAATTGCAGTCCGAGGAAGGCTAACTGGAATC
 AACATACCCGTAGCTTTAGAAAGCAGTTTCCGCACCAGCGAAGAGTACAAGAGCGATGGA
 ACCCATGTTCTGGAAGTTTGCACATCAGAGTAAACAACTTGAAAACCCCTCTTGATA

FIGURE 23.1

55/64

GCAGAATTCACCCAGCCTTGTTCCATTTTCTCTTAACAAAACACACCCGCAAAAGCTCTCA
CAAGCTGCTTTTGATGAAGCCACATGTATTTCCCCCTTCACAATTTACAGGAAGTTACTCT
TAAAAGAAAGTGATTCTGGTGTTTACCGCCTGTGTTAAAGGGACAGAGTTCCTTTTATT
TCTGATAACGTTTGAGCGAAATACAGAACTATCTGTAGACTAGCATAGTCGGTACGTGA
GTAAGGAAAAGCAATAACCTGCTGTCCGGTGAGCACAAAATTCCTGCTACGAACAGTGCC
TTACTGCTGCTTGAGACTGCAAGTCGCAGATCACACTAGGTATTGACTGATTGTATAAG
GAAATTTCTTAAAGTCTAAAGTAAAGGTGGTACCTCCTAAAAAGAGGGGAAGAGAGAAAA
CTTTGTGTGGAAGGATAAGGAGTGTGTTTATAGTTTCAGTAAGAGTGTACGTTTAAATTT
TTCTTCTTCTCTGCCTCTTTGCCAAGTAGCCTGAGTGCATCTGTTATCCAGAAGTAGTA
TTACTCTAGGACAAACTTCAAATTCCTTATTCTGCGTTGCCTTTAAGGAACAACATACTT
TCTTCTGTTCTTTTTTCCAAAAACACACGCCTATGGCTCTGTGTGTGGTGTTTAGCCAG
CCTCCTCCCAGATAAGGGGTTCCCTTCCCTCCTTTGCATTGAAAGGAAAGTGCAAGTCTG
GACATGTTTATCAAGAGGAAAAGTGACTTCTCAGTAATAGACTGTCAAATTCGGGCTGCT
GCCCCAGTGTTCGCTTTGTTATGGCAGGTGAAGTTCACCTTTGCCCCACCCAGTGTTCCT
ACAAAAAGGCAAGGTTCCAAGTATTCATATGAACAAGTGTTACTTTAGGACTTGAGAGGT
TGGGGGTGGAGGATGTTTGCATAGTTGAAGCCTTGGGCGGGGGTGTAGGAAACGGCGAGT
ACAGAGGCCATAGAAAAAGCTAAGACTCAGTTTGACGTCGTGAGCCGGCTTGGTCTTCTA
CCCAGTGACTCAAAGCACTAAAAGTCAGCATAATCGGAACTGAAGTCAGTAGCATCGCCC
ATTTGCCATTCACTGCAGTAGCAAAAGTAGTACTCTGTGGTGGGTTAATCGGTTTGAGGC
AGCTCCTTAAATGAACATTTGTGTTTCATTTTCTGTTATTTTCCCGAACATGAAAAGAC
GATAAACTGAAATGGAAAAGGTAAGTACAAAAGTGTGCCTTACCTGTTTCCGCCCTGA
TTTCTGCTGATTCAAGACTATTCTGGCTAAACTGATTGGATTCTTTTTCTAACTAGGCAG
TAGGGGATCAGAAATCACACACGGTACCGGCTGTGTTTATTCTGAGAGGTGCTGGGGAGC
TTTGGGTCTGACTTCCTTTTACATGCCTGTCTTCTCTTTTGGACAGATCTATTCCAGAGG
GGAGCTTCACCACCTTCATTGACGGCTTTAATGAAGAGAAAAGCAACTGGATGCGCTATGT
GAATCCAGCACACTCTCCCCGGGAGCAAAACCTGGCTGCGTGTGAGAACGGGATGAACAT
CTACTTCTACACCATTAAAGCCCATCCCTGCCAACCCAGGAACCTTCTGTGTGGTATTGTG
GGACTTTGCAGAAAGGCTTCACTACCCTTATCCCGGAGAGCTGACAATGATGAATCTCAG
TAAGTGGATTACAGAACAAAAAATAAAAAATGCCAGTAATGTCGGTTCTGCCCCTTTGA
ACTAATAACATGTTGTTTAAATTATACGGCTTTGTGATGTGTTGGATGAAGTAGGTGGCTT
AAGCTAGGGACTAGGAAGAGGAAAAACATTTTTTTGAGTCCCTATTAAGTATTAGGAACT
TGATCATTTTAAAGTATATATATATATAGAGGAGCTACCTTGAGTTTTGAATTCAGGATGT
TACAGGAAGAAATATATGTCCAATTCTAATTTATCCAAAAGCAGTTGGGAGAATTACAGG
GATTGGTCCAGACATGCTGCGTATGCAAGGTATAGCCCTCATCTGTGGTACTTTGGCAGG
GCTTAGACTGCATCAAAATATTTATAGATGTACATTTGAGTGTACAGTTAGGATCTGATG
TGGAACATTGTAAGATCATTGCTAGAAAACTTTGTGATAATTTTTCAATATTATTCTAA
GTGAATAACCGTAAAGATTTTACATCTTAGCTTCCTTCCTTACAGTAAAAAACTATCTG
ATCTCTTGATCAGTATTATAGTAGCCACCTATCACTTTATCTTAACAAATTCTCAATTCC
TTAGGTTTATGTGCTTTTACTTCTTTTATTTGATTAAAATTGCTGTGATGACCTCTCTCT
GCAGAGGGCTGCATCATTTTGGTCATTCTCAAGTGATCTCTTTGAGCAATTTAAGAATTG
CCATAAGATTCTAACCTCTGCTGTAACCTATGGTTGTGTGTTCTTGGTTAGACCACTAAAT
CTTATTAGCAGTTTTTAAAAATTATTCTTTTGGTTTTAGAAGTTAAGACTAAATGCTGAAG
TTTTTGTAACTTTTGGTTTTGATATCATTTCAAACCTTAAGAAAACATTTGAAGAAAAGGA
CAAAGAATTTCCACTTACCCTTTACCAGGTTTACCAGTTATTGATAAGTATATCCATTT
GCTTTACCAGAAGGCTAACTTGTTTTAGTTCTCATTTTCACCTTTGAGACATTTGGAATA
AATATCAATGTAAACATAAATTGGAATTTGACTTTGATTTTAGGACCAATGAACAAGCC
AAGTACTTACCCTAGTCATATATAATCCAAGTGTATGGTTATTTGGTATTCAATCCACAC
TTCATTTTACTTGATCTCCCTTAAGATTGCAAGATTGTGTTTGCAGTTTTTCTGAAAATC
TGGGGCTATAAAAGCATCAGGACCTCCCCGTAGGGGAGGTCGTGTGTTTGGGGTCCTTA
CACAACAGGTTACCCTTGAGCTTCAGGAAAAGAACTGGCTCTCAGTTCCCCAGTTCCAGC
TTAATGGGTCTAATTAGGTCTGACCAAAAAGGTGGCAGTTCTTTTCCCTCATGTCTCTT
CAGCGCTCCCCGAGACTCTGGAGACTCTGTGATATCCCTAGGGCTGAGCCTCCAGGAAC
CATTCGGCTGTTGTGGCATCTGTGTATGCCATGCCAGTGCTGAGGACCTAGTAACAAAC

FIGURE 23.2

56/64

GACAAATGCACAGGCACAGTGGCATT TTTTGTGGAAC TCGTATTCAGCTGTGCGTCTCAG
AAGAAGCGCACAGCTCCCTCCTGGCTTTCTTAACATAGTGAGCCACTTCCACTTAAGGGT
CTCCTTACATTCCCTTGAGTTTAATCATTTCATGGATTTCAGAGGAAAGTCTTTTGATTTTTG
CTTTTCTTTAAACAGTTCATTTGAGGTGACCTACCCAGTGACTTTGCACCAACCACCAA
GAAACTTTTTTGCATGCTTCCCGCACCTGTGCCAATCAAGGGAAGGGTTTAAAGGCCTG
GCGTTTTTATTCTCAAAGAAAGGTTTTGTCACAGTATTTTAAGGTTCAAGTGCTTCTACT
TTGTGTTTCAAGCAACTGTCATATATACTGTGAAATGACACCTTTTATTTATCCCTTTT
TATTTATGCAGTATGTCCCTTTTATTTTGGCAGAATTTTTTCTAAATGGTGGTTTAAACA
TTTTCAAGCACATTTTCATTGTCCAATATTCATAGTAAAGAATGAGAGTTAACAATAACCA
GTCACATTAACAAGATTCCCTGCTGCCAGTTGTGAAACCGGTTGTCTTAGGCGTGGCAG
CTGATGATTGAGACTGTGATCAGGAAAATTTCCACTATTTTCATCAGGCCTAATAGGTAGA
TTGTGTCTCCAAATGAACTGTGTTGGGTTTCCATGCTTAAAGCACATAGAGGTGGTGCA
AGAATCTCCATGAGGGCTTAAATGGCAGTGATGGTTTTCAGGCGGTAGAGTTTGGAGAAGAA
GGGATTTGAAACAAACCAAGGAAAGAAAGTAAGTAGCCAGAAATCACAAAATGGCATT
TTTCTAAAAACAAAGGAAAGGAATAAAAGAACTAATAAGTTTGAAACCCCTACCCCTCC
CAAATTTGGCAGGGGGGAGGTATTTTTTTTCTATCTATCTAACTAACCCATCTAGAAAA
CAGTTGACCAAATTATAGACTTCTAAATGTTAATCTGCTTTCTCAGTTTCAGTTGAAAAAG
AGACTTTGTTTTGCCTACTGCAGAACTTCTAGGTTCTTTCTTATAGTCTTGGGGTTCTTA
TTATAGATCGAAAATGTGAGTCGGCATAATTAAGCCATTTCGGAGTCTTCAGAAGCAGTTC
ACTCTTGAAATGACTCCGTCCGCCTACAGCCATTTAAGATTTTCAGAACAAAAACAGATCT
TGATTTTCTTTTTTCATGTTAACTCAAGCTGTTGCTGAGTGGGAGAGTCAGAAATGACACC
AGCTCCACTGATTACTCAGCTGCTGAAGGATGATTTTTTAAAAATGCACCTTTACTGTATA
TGGACTTCCTAATTTCCACCTGTAGAGCATCTTAGGGAGGCTAACATGTCACTCTGGATG
TTCTTTTGAATAAGATGCAAATCTATTTTTCTGAAGGCATTAGAGATAGCAAACATTTA
TTGTGAGTTTACTATATACTAGGCACTGTGCTAAGTGTTTTGCATAGAAAAGTTTTAAAT
CTGGCTTTTTTGTGGCCCAATCATAAGTTTCATATCAGTTCAACATTCAAATTATATTA
AGGTACTTAAGAAGAATCCCTGGCTAAATGTGAGGGCAGTGCCACAGATGGACTGAAAC
TTTATGCTTATTGCACATTTATGCTATTATTATTGTTGAATTATAGAACCAAGGGAGTG
TGGAAGCCACTGGAAAAAATATGAGACTTAGATACATAATTTGAGTAAAAATGGCTCAAA
GTCATGAGGGTAAAGTTTTTTGTATTTCCATTTTATTCGAGCGGCATCGTTTTTAAAAAT
CATTATGAATTTGACCCTATATAGATGTTTCCAAATAATTCTTTTTACCTTCATAAAAT
TCCTTCCTGTGGCTGTGAGATGCCTTGCCATCAGTTTTCAAGCTTAGTTGTCTTTCTCA
CCTCACTGAAAGACCGATATTGGCCTGATAAGGAGATATTTATTTTGTGTTTAGTGCTTC
AGAAATCCCTCTCCCTCAGCAAGCTTTCCATCACGGCCCCCGTCAGCATCTTCCCTGA
TAGCGTTCTTCTCTGTGTTTATTCTGGGGCTTCAGGCTCGCCCAGGAGGAAGTGAACCC
GCTGGCAGGAGATAACATTCTCTAAGGGGCTCTCAAATTGGAATCGAATCCCTCAAGCCA
GTCAGCCTAGAGAATACATTTAAAGGGTTCAGTTCTGGAGTTTCACAGAGTTTCATTTCTA
GACCTATCAGATAGCAAGTGTGGAGTTCTTTCTCAACTAAATTCAAGCAGAGACATTTT
TAGACGATGAAGGATATTTGCACAAAGGCTTCAGCATGATCCCCCAAACCTGCTGCCTCT
GAAGGCATCTCCACACATTGACAGCCAATGCCTTCAGTGCGTTCTTAGGGCAGGTGTCCT
GGCTTGAGTGACTGTCTCCAATAATCAGAGCTCAAACTAAACATCGTATGTTTTACTTT
TGGTTTCCAGGCAAGGCTGAGCAGGGAATTTTTCAGTTTTCCCTGCCAGATGGGTGTTTT
TTCCTGAAGGCATCATTTATTGTGTAGCGAGGAGACAGGGCTGGCTGTGGCAGGGATAGT
CTAGAAGTGTCTCATTTGCTGCTGTTTCTTAAATAGTATCTTTACCAAGTAATAACGTGCC
GTCTTTGGGAATAAGTGCTTTCTCTTAGCTGTTCTGTTTTCTTGGGTGCGCTAAGTAA
TTGAACTGGCTCAGGAAGTACCTATTGTGTTTTGGCAGAGGTGACTGTCACGCCTTGTTGA
CTCCAGGGGCCAGCACTGCTGGGATCCTGGCTAGACCAGACAGAGCCTTGGTGAAGTGCT
TAGGCTGTCTGCACATCGCGAGGAAGGTGGTATTCACTTCGCTAAGCTCCTTGGCATAGG
CAGTTTGAACAGGGCTTTATCAAATTCGTATTCAACAAGAGTAGAAGCGAAAATTGATGA
CTGTGTATTACTTGAAATGAGTCTTAATCTTTACATTTAGTTCTCAGGGTATGCTGATT
TCCTTTAGGTAAACCATGAACATCAGAAAGACTTTTATTAACCTATGACAGGGTCCCCAC

FIGURE 23.3

57/64

CCCAGTATTTTCCACTCCATTAAATGGAAGTTTTTTTTTTTTTTCTTTTTTGAGAC
 AGAGTTTTGCTCTTGTTGCCAGTCTGGAGTGCAATGGCACAATCTCGGCTCACCACAAC
 CTCCACCTCCCAGATTCAAGCGATTCTTCTGCCTCAGCCTCCCAAGTAGCTGGGATTACA
 GGTGTGCGCCACCACGCCCAGCTAATTTGTATTTTGTAGTAGAGATGGGGTTTCTCCATG
 TTGGTCAGGCTGGTCTCGAACTTCCGACCTCAGGTGATCCGCCCACCTCGGCCTCCCAA
 GTGCTGGGATTACAGGCAAGAGCCACTGCATCCAGCTTAGGCTATCTTACTCCAGCCTAA
 ACAGCAATTTTCTATCATAAGGTCTGTACTAATGAAAACAGAATCACCCAAGGCTGCTGT
 TTGTTCTGTCTGTGCTGCCATTGTCCGCATTTTGTCTGAGGAGGAAACGGAACCTGCACTTT
 TGAGTGAGTGGCCAGAGCCTTCTAGAATGAGAGTGCGTTGGAAGCCAGATATGTGGCGA
 TTGTGTGCGCCAGCTGTTACTCAGGTTTTCTCAAGAAGGAGGAGCAACTTTGGCAGTTTTG
 CTTCAGTTCTCTCTAGCCCTCTGTGTAATCGCCCCTTTTCTTTATTTTCAGCACAACAC
AGAGCAGTCTAAAGCAACCGAGCACTGAGAAAAATGAACTCTGCCCAAAGAATGTCCCAA
AGAGAGAGTACAGCGTGAAAGAAATCCTAAAAATTGGACTCCCAACCCCTCCAAAGGAAAGG
ACCTCTACCGTTCTAACATTTACCCCTCACATCAGAAAAGGACCTCGATGACTTTAGAA
GACGTGGGAGCCCCGAAATGCCCTTCTACCCCTCGGGTCGTTTACCCCATCCGGGCCCTC
TGCCAGAAGACTTTTTGAAAGCTTCCCTGGCCTACGGGATCGAGAGACCCACGTACATCA
CTCGCTCCCCATCCATCCTCCACCACTCCAAGCCCCTCTGCAAGAAGCAGCCCCGACC
AAAGCTCAAGAGCTCCAGCCCTCACAGCAGCCCTGGGAATACGGTGTCCCTGTGGGCC
CCGGCTCTCAAGAGCACCGGGACTCCTACGCTTACTTGAACGCGTCTACGGCACGGAAAG
GTTTGGGCTCCTACCCCTGGCTACGCACCCCTGCCCCACCTCCCGCCAGCTTTCATCCCT
CGTACAACGCTCACTACCCCAAGTTCCCTCTTGCCCCCTACGGCATGAATTGTAATGGCC
TGAGCGCTGTGAGCAGCATGAATGGCATCAACAACCTTTGGCCTCTTCCCGAGGCTGTGCC
CTGTCTACAGCAATCTCCTCGGTGGGGGCAGCCTGCCCCACCCCATGCTCAACCCCACTT
CTCTCCCGAGCTCGCTGCCCTCAGATGGAGCCCGAGGTTGCTCCAGCCGAGCATCCCA
GGGAGGTGCTTGTCCCGGCGCCCCACAGTGCCTTCTCCTTTACCGGGGCCCGCCAGCA
TGAAGGACAAGGCCTGTAGCCCCACAAGCGGTCTCCACGGCGGGAACAGCCGCCACGG
CAGAACATGTGGTGCAGCCCCAAAGCTACCTCAGCAGCGATGGCAGCCCCAGCAGCGACG
AAGCCATGAATCTCATTAAAAACAAAAGAAACATGACCGGCTACAAGACCTTCCCTACC
CGCTGAAGAAGCAGAACGGCAAGATCAAGTACGAATGCAACGTTTGCGCCAAGACTTTCCG
GCCAGCTCTCCAATCTGAAGGTAGGCCTTGAGAGAGAGCAGTCCAAGGGGCTGTGAGTGC
 ATGCTTGTGTTTGTATTTAGCTTGCTTTCCATGGGGTATCGATTGCATTTGCAGTAGTAT
 GAGCCCCCGGTTGGGGATAGTGGGTATGGATTCCGCCCTGGCTTTTGCCACTTCTAGCTCT
 TTGACTTTGGACAAGTGACTTCCCTTCTCCTGATTTTCTTCTGAATAATAAAAAAATTAG
 GGGTTTGGACTAGAAGATTAGGTGAACTCCCTGCTAGCCTGTGATTTTGTGCTTTTAA
 GAAAAACACCATTCTGAAAACATGAAGATTTCTTCTTTTAAAGACTGTCTTGATGCTTTT
 CTTAAGATATTTGCATCAACACTTGAGTCTTGAGAGCAGAAATGTTAGGTCTCAGAGCAG
 CTTGAGAGCAGAGCTAACACATGTGGCTTCTTCCAGGTCCACCTGAGAGTGACACAGTGG
AGAACGGCCTTTCAAATGTCAGACTTGCAACAAGGGCTTTACTCAGCTCGCCCCACCTGCA
GAAACACTACCTGGTACACACGGGAGAAAAGCCACATGAATGCCAGGTGCGCAGTATTTT
 CTGGGTAGACCTTCTGACCTTTGTAGAAAATGTCTGTGAGTCACCTCCCATGTCTCTATA
 TAGCCCGTAGTTAAAGCCAAACACCAGATTCTGCGTTGTCCCATCTGGACTGATGGCACT
 ATGGTCCCTCCAGTACTTTGTATCTGCTGATGACTTGAGATGGCACAGCCAGCTTCCAG
 TGGGTGGGAAAATGGTAGGGGAAATAAACAGCCCCCTCGTGTGCTGTGTGCCACATCCCC
 CCGTTTGTCTAATACCACACTGGAGGTGCCACAAGGAGGCTTCTCACCTCCTAGGTTGCT
 GGGCGTTGGCCGGTAAGCCTGCCCCCTCCCGTTGGCAACTCTTAATCTTCTGGCCTTCCCTG
 TCTCCCTTCCCTGCTGTCTCTCTCCCCTACACTGTAGGCTGCGCACAGAGATTTAGCAG
CACCAGCAATCTCAAGACCCACCTGCGACTCCATTCTGGAGAGAAACCATACCAATGCAA
GGTGTGCCCTGCCAAGTTTACCCAGTTTTGTGCACCTGAAACTGCACAAGCGTCTGCACAC
CCGGGAGCGGCCCCACAAGTGCTCCAGTGCCACAAGAACTACATCCATCTCTGTAGCCT
CAAGGTTTCACTGAAAGGGAAGTGGCTGCGGCCCCGGCGCCTGGGCTGCCCTTGGAAGA
TCTGACCCGAATCAATGAAGAAATCGAGAAGTTTGACATCAGTGACAATGCTGACCGGCT
CGAGGACGTGGAGGATGACATCAGTGTGATCTCTGTAGTGGAGAAGGAAATCTGGCCGT

FIGURE 23.4

58/64

GGTGAGAAAAGAGAAAGAAGAACTGGCCTGAAAGTGTCTTTGCAAAGAAACATGGGGGAA
 TGGACTCCTCTCCTCAGGGTGCAGCCTTTATGAGTCATCAGATCTACCCCTCATGAAGTT
 GCCTCCCAGCAACCCACTACCTCTGGTACCTGTAAAGGTCAAACAAGAAACAGTTGAACC
 AATGGATCCTTAAGATTTTCAGAAAACACTTATTTTGTCTTAAAGTTATGACTTGGTGA
 GTCAGGGTGCCTGTAGGAAGTGGCTGTACATAATCCCAGCTCTGCAAAGCTCTCTCGAC
 AGCAAATGGTTTTCCCCTCACCTCTGGAATTAAAGAAGGAAGTCCAAAGTTACTGAAATCT
 CAGGGCATGAACAAGGCAAAGGCCATATATATATATATATATATCTGTATACATATTA
 TATATACTTATTTACACCTGTGTCTATATATTTGCCCCCTGTGTATTTTGAATATTTGTGT
 GGACATGTTTGCATAGCCTTCCCATTACTAAGACTATTACCTAGTCATAATTATTTTTTC
 AATGATAATCCTTCATAATTTATTATACAATTTATCATTCAGAAAGCAATAATTAAAAAA
 GTTTACAATGACTGGAAGATTTCCTTGTAATTTGAGTATAAATGTATTTTGTCTTGTGG
 CCATTCTTTGTAGATAATTTCTGCACATCTGTATAAGTACCTAAGATTTAGTTAAACAAA
 TATATGACTTTCAGTCAACCTCTCTCTCTAATAATGGTTTGAAAATGAGGTTTGGGTAATT
 GCCAATGTTGGACAGTTGATGTGTTTATTCCTGGGATCCTATCATTTGAACAGCATTGTA
 CATAACTTGGGGGTATGTGTGCAGGATTACCCAAGAATAACTTAAGTAGAAGAAACAAGA
 AAGGGAATCTTGTATATTTTGTGTATAGTTTCATGTTTTTCCCCCAGCCACAATTTTACC
 GGAAGGGTGACAGGAAGGCTTTACCAACCTGTCTCTCCCTCCAAAAGAGCAGAATCCTCC
 CACCGCCCTGCCCTCCCCACCGAGTCTGTGGCCATTGAGAGCGGCCACATGACTTTTGC
 ATCCATTGTATTATCAGAAAATGTGAAGAAGAAAAAATGCCATGTTTTAAACCCTGTC
 GAAAATTTCCCCAAAGCATAGGTGGCTTTGTGTGTGTGCGATTGGGGGGCTTGAGTCTGG
 GTGGTGTTTTGTGTGTGGTTTTTGTGTCTTTTTTTTTTTTTTTTTTTTTTAATGTCAAAT
 TGCACAAACATGGTGTCTTACCAGGAAGGATTCGAGGTAGATAGGCTCAGGCCACACTTT
 AAAACAAACACACAAACAACAAAAACGGGTATTCTAGTCATCTTGGGGTAAAAGCGGG
 TAATGAACATTCCCTATCCCCAACACATCAATTGTATTTTTTCTGTAAAACCTCAGATTTTC
 CTCAGTATTTGTGTTTTTACATTTTATGGTTAATTTAATGGAAGATGAAAGGGCATTGCA
 AAGTTGTTCAACAACAGTTACCTCATTGAGTGTGTCCAGTAGTGCAGGAAATGATGTCTT
 ATCTAATGATTTGCTTCTCTAGAGGAGAAACCGAGTAAATGTGCTCCAGCAAGATAGACT
 TTGTGTTATTCTATCTTTTATTCTGCTAAGCCCAAAGATTACATGTTGGTGTTCAAAGTG
 TAGCAAAAAATGATGTATATTTATAAATCTATTTATACCACTATATCATATGTATATATA
 TTTATAACCACTTAAATTGTGAGCCAAGCCATGTAAAAGATCTACTTTTTCTAAGGGCAA
 AAAAAAAAAAAAAAAAAAAGAACTCCTTTCTGAGACTTTGCTTAATACTTGGTGACC
 TCACAATCACGTGCGGTATGATTGGGCACCCCTGCCTACTGTAAGAGACCCCTAAAACCTTG
 GTGCAGTGGTGGGGACCACAAAACAACCAGGGAGGAAGAGATACATCATTTTTTTAGTATT
 AAGGACCATCTAAGACAGCTCTATTTTTTTTTTTGCCACTTTATGATTATGTGGTCACACC
 CAAGTCACAGAAATAAAAAACTGACTTTACCGCTGCAATTTTTCTGTTTTCTCCTTACT
 AAATACTGATACATTACTCCAATCTATTTTATAAATTATATTTGACATTTTGTTCACATCA
 ACTAATGTTACCTGTAGAAGAGAAACAAATTTGAAATAATCCAGGGAAACCCAGAGCCT
 TACTGGTCTTCTGTAACCTCCAAGACTGACAGCTTTTATGTATCAGTGTGTGATAAACA
 CAGTCCTTAACTGAAGGTAAACCAAAGCATCACGTTGACATTAGACCAAATACTTTTGAT
 TCCCACTACTCGTTGTCTTTTCTCCTTTTGTGCTTTCCCATAGTGAGAATTTTTAT
 AAAGACTTCTTGCTTCTCTCACCATCCATCCTTCTCTTTTCTGCCTCTTACATGTGAATG
 TTGAGCCCAACAATCAACAGTGGTTTTATTTTTTCTCTACTCAAAGTTAAACTGACCAA

FIGURE 23.5

FIGURE 24

60/64

GAATTCCGGGAAGCCAGACGGTTAACACAGACAAAGTGCTGCCGTGACACTCGGCCCTCCAGTGTTGCGG
AGAGGCAAGAGCAGCGACCGCGCACCTGTCCGCGGAGCTGGGACGCGCGCCCGGGCGGCCGACGAAG
CGAGGAGGGACCGCCGAGGCTGCCCCAAGTGTAACCTCAGCACTGTGAGGTTTCAGGGATTGGCAGAGG
GGACCAAGGGGACATGAAAATGGACATGGAGGATGCGGATATGACTCTGTGGACAGAGGCTGAGTTTGAA
GAGAAGTGATACATACATTGTGAACGACCACCCCTGGGATTCTGGTGCTGATGGCGGTACTTCGGTTTCAGG
CGGAGGCATCCTTACCAAGGAATCTGCTTTTCAAGTATGCCACCAACAGTGAAGAGGTTATTGGAGTGAT
GAGTAAAGAATACATACCAAGGGGCACACGTTTTGGACCCCTAATAGGTGAAATCTACACCAATGACACA
GTTCCCTAAGAACGCCAACAGGAAATATTTTTGGAGGATCTATTCCAGAGGGGAGCTTCACCATTTCATTG
ACGGCTTTAATGAAGAGAAAAGCAACTGGATGCGCTATGTGAATCCAGCACACTCTCCCCGGGAGCAAAA
CCTGGCTGCGTGTCAGAACGGGATGAACATCTACTTCTACACCATTAAAGCCCATCCCTGCCAACCCAGGAA
CTTCTTGTGTGGTATTGTGCGGACTTTGCAGAAAGGCTTCACTACCCTTATCCCGGAGAGCTGACAAATGA
TGAATCTCACACAAACACAGAGCAGTCTAAAGCAACCGAGCACTGAGAAAAATGAACTCTGCCCAAAGAA
TGTCCCAAAGAGAGAGTACAGCGTGAAAGAAATCCTAAATTTGGACTCCAACCCCTCCAAAGGAAAGGAC
CTCTACCGTTCTAACATTTCACCCCTCACATCAGAAAAGGACCTCGATGACTTTAGAAGACGTGGGAGCC
CGCAATGCCCTTCTACCCCTCGGGTCGTTTACCCCATCCGGGCCCCCTCTGCCAGAAGACTTTTGAAGAC
TTCCCTGGCCTACGGGATCGAGAGACCCACGTACATCACTCGCTCCCCCATTCATCCTCCACCACTCCA
AGCCCCCTCTGCAAGAAGCAGCCCCGACCAAGCCTCAAGAGCTCCAGCCCTCACAGCAGCCCTGGGAATA
CGGTGTCCCCTGTGGGCCCCGGCTCTCAAGAGCACCGGGACTCCTACGCTTACTTGAACCGCTCCTACGG
CACGGAAGGTTTGGGCTCCTACCCTGGCTACGCACCCCTGCCCCACCTCCCGCCAGCTTTCATCCCCCTCG
TACAACGCTCACTACCCCAAGTTCTCTTGCCCCCTACGGCATGAATTGTAATGGCCTGAGCGCTGTGA
GCAGCATGAATGGCATCAACAACCTTGGCCTCTTCCCGAGGCTGTGCCCTGTCTACAGCAATCTCCTCGG
TGGGGGCAGCTGCCCAACCCATGCTCAACCCCACTTCTCTCCCGAGCTCGCTGCCCTCAGATGGAGCC
CGGAGGTTGCTCCAGCCGGAGCATCCCAGGGAGGTGCTTGTCCCGGCGCCCCACAGTGCCTTCTCCTTTA
CCGGGGCCGCCGCCAGCATGAAGGACAAGGCCTGTAGCCCCACAAGCGGGTCTCCACGGCGGGAACAGC
CGCCACGGCAGAACATGTGGTGCAGCCCAAGCTACCTCAGCAGCGATGGCAGCCCCAGCAGCGACGAA
GCCATGAATCTCATTAAAAACAAAAGAAACATGACCGGCTACAAGACCCCTTCCCTACCCGCTGAAGAAGC
AGAACGGCAAGATCAAGTACGAATGCAACGTTTGCGCCAAGACTTTCGGCCAGCTCTCCAATCTGAAGGT
CCACCTGAGAGTGCACAGTGGAGAACGGCCTTTCAAATGTGAGACTTGCAACAAGGGCTTTACTCAGCTC
GCCCACCTGCAGAAACACTACCTGGTACACACGGGAGAAAAGCCACATGAATGCCAGGTCTGCCACAAGA
GATTTAGCAGCACCAGCAATCTCAAGACCCACCTGCGACTCCATTCTGGAGAGAAAACCATAACCAATGCAA
GGTGTGCCCTGCCAAGTTACCCAGTTTGTGCACCTGAAACTGCACAAGCGTCTGCACACCCGGGAGCGG
CCCCACAAGTGCTCCAGTGCCACAAGAATAACATCCATCTCTGTAGCCTCAAGGTTACCTGAAAGGGA
ACTGCGCTGCGGCCCCGGCGCCTGGGCTGCCCTTGGAAGATCTGACCCGAATCAATGAAGAAATCGAGAA
GTTTGACATCAGTGACAATGCTGACCGGCTCGAGGACGTGGAGGATGACATCAGTGTGATCTCTGTAGTG
GAGAAGGAAATTCTGGCCGTGGTCAGAAAAGAGAAAGAAGAAACTGGCCTGAAAGTGTCTTTGCAAAAGAA
ACATGGGGAATGGACTCCTCTCCTCAGGGTGCAGCCTTTATGAGTCATCAGATCTACCCCTCATGAAGTT
GCCTCCAGCAACCCACTACCTCTGGTACCTGTAAAGGTCAAACAAGAAACAGTTGAACCAATGGATCCT
TAAGATTTTCAGAAAACACTTATT

FIGURE 25

[illegible]

FIGURE 26.1

62/64

GCTACTGCCACCGCCACGGCCACCACCACAACTACTACCACTACCATTTCCACCATCACCTCTACCATCA
CTACTGGCCTCATGGATAGCAGTCACCTGGAGATGACGTCTGGGCGGCTCTGCCCCTTCTATCCAGCAG
CAGCACTAATGTCCGGAGACCCAAGCTCACTTTTGATGACTCGGTTCACAATGCTGATTATTACATGCA
GAAGCTAAGAAGCTGAAGCACAAAGCTGATGCACTGTTTCGAGAAATTTGGCAAAGCTGTGAATTATGCTG
ATGCCGCCCTCTCCTTCACTGAATGTGGCAATGCCATGGAACGCGACCCTCTGGAAGCAAAGTCCCCATA
CACCATGTACTCTGAGACTGTGGAGCTCCTCAGGTATGCAATGAGGCTGAAGAACTTTGCAAGTCCCTTG
GCTTCGGATGGGGACAAAAAGCTAGCAGTACTATGCTACCGATGTTTATCACTCCTCTATTTGAGAATGT
TTAAGCTGAAGAAGGACCATGCTATGAAGTACTCCAGATCACTGATGGAATATTTTAAGCAAATGCTTC
AAAAGTCGCACAGATACCCTCTCCATGGGTAAGCAATGGAAAAGAACTCCATCCCCAGTGTCTCTCAAC
AACGTCTCCCCCATCAACGCAATGGGGAACTGTAACAATGGCCCAGTCACCATTCCCCAGCGCATTACCC
ACATGGCTGCCAGCCACGTCAACATCACTAGCAATGTGTTACGGGGCTATGAACACTGGGATATGGCCGA
CAAACTGACAAGAGAAAAACAAAGAATTCTTTGGTGATCTGGACACGCTGATGGGGCCTCTGACCCAGCAC
AGCAGCATGACCAATCTTGTCCGCTACGTTCCGCAAGGACTGTGTTGGCTGCGCATCGATGCCCACTGT
TGTAGTGGGTGTTCTCAGATCTCTAGCATCAGACCATCACTCTACCTCTACCAGCGCACTGATGGTCA
TGGTGGAACCTCACTCACTGGGGAACGTTCTCTTTGGTTATGTTTGTGTTTTATGCTTCTTTTGTTATCT
GTAAAAAACAGAAGTCATTGTAAGTTGACACTACAACCTTAAGGGCAGTGACGTTTTATTACTTAGTCAT
TTTTTTTCTTTTAGCATTTTGATATGCATTTTCTCAGATTCCACCATCTTTTGTGCTTTTATGGAATGACAG
TCCCTACAATATTGTTTTAAGCCACACTACCCAAAACAAAGAAATGGGAAGCACTTGTGATAAAGACAGG
CTCCTGAGAAATGCAACAAGTGGTCTTACATATACATGAGAACTTAGACACAAGGGACCATCCCCAAAC
TCTACTCTTATACCCAGAAAAAGACATATTTCAGAATCTGTCAAACTTTTGTGTATCCACAGATTCAAT
CTTCAGGTGAGAATTTTCATTGTCAAAACCCACTGGTTAGATGTTGTAGCAACATCATAAAATCAAGAGT
ATCAAGAAAAATAATGAGCATAGCAATGCTACTCTTAAAAAGATGCTATGCCACACAACCAGAGGACTTT
CTTGTTAGCATCCCTTTCTGATTCCCTATTTTGTTAATTTAATGATAAGAAGAAAGGGTGACATTTAT
TTTGACAAGTTTTAGGCATCAGCTGGCATCAGTGTTTTTCAACTCCATTATTTGAAGTGTAATCCTCAC
CTGGGGTTCTCTGTGTGCAAAGCTGTCTTTTGAAAGACAGTTTGGTTGATGCATGCCTTAGTAGCCAAA
ATGCTACACTCTAGACTTACAAGTGGGAGTTAAGAGAGGTCTGGAAAGTGTTCAACAAGGAATTCACACC
TCTGCCCTCCTTTGCAACAACAACATTTTACCAGTTGGTAAGTGGGTCCATAACTGGCAGGATTTTTAAAT
TGTATTTTGCTCAAAATCTATGGGAACAAAAGTCAAGGTATCACTACCTAGAAAGTAATGATATACAGTTTT
CTTCCTAGTGCGCTTGAAAAATCTGGACTTCCTCAATTATTATTACATTTTTCTCTCTTATAGGTTTTCTGT
TTTCTACTTTCTTTTTTCTCTTATCTGTGTTTCCCTTTCTCTTTGTTGGCTCATTAACTTTTGACTGAAT
TACAATTACTCCTTTTATTAAAGTCCATATTTGTGAATCATTTCCATGAAAAATTTCTAAGAAAACTCC
AAACTCTCTAAATAGTAGCTAACTTTTATTTTTTTAAAAATGAGTCGTGGGGTAGTGCTTCACCTTGAGAT
GCTTTGAAAGAGCCCTAAACATTGGGAACCATTACCTAATTTGGAGACATTTCTCACTGGTTGTGACTA
CCCCCTTATGATCCTTACATTCATTTTATGTCTCCCTAAACATCACAATGTAAATATCATTTTGTATGTTT
CAGCTCACCAGAAGATTCTTACACTTGGGGTAAACATATCCATGCATTACTTACTGTTAATTACCTGC
GGTATATAAATTCATGTAGCCTTTAATATGCTGGGTTATCAAATTTCTGTTCACTGAGTTATGACCAGATA
AATAATAGATATGCACATGAAAGATGCAAACTTGTGTGATTATTAAAGCCAGCCATGCAGGTCCATGATA
GAAACAGCAGGTGATGACTCTGCACTCTCATGTCAAGGTTAGCTATATCCCCAGTTGCAAAACAGCCAG
ACTTGAGCTGTGCTCTGGTCATCTTTGAGTTTAAAGGCTTTTGTGTATAAGGCTGTGGAAGTTGTACTC
CAATGGCTGAAGCCATGTTGTTAATATGGCTGATGGGAGCATCCCTGCAGCTGAACCCAGCACTTTTTAT
GCTCCCACTGTGGTTGAGCTTTATGTTTACAGTCTCAGCAACAACACTTATGCATCCAAACACTCACAA
TGAAACCTGAAAGAATCTTTCTGAGCCTCTTAAAGAGAGAAAATGATGATAACATTAAAGACTCTGAAC
ACCCAAGGTTGGTGTCACATATAAAAAATTAAAGCTGATGACTTTGCAGTGACTCAAGTTGTCTCTTTATCA
TGGTTTACCAGGTAGAGTGCTGGCTATTACTATATAATGAAGCCCACTGGCTTGACTTGTAAGTTCAAC
CTAAACCACAATCCTAGACCATCATGGATTTAGGAGTAGATTCTTCTTGAAATCCCACATCCAGAACTA
GACATTAGAATGTTGAGGCAGTTTCCCAGAGAAAACAAGCATATTGCCTCATGGATGAAAGACTTGTAGTT
CTAGTTTCAGTGACTTGTTATATCTACTTACATACAACAGGGAGGCAAGAGGATTCTCTGTCTATCTCTGG
TGACTGAGTGTAATAATATGTGCCAAGTCTGCAGCACAGTGACCAAACTGACAATCGAGCTCTGGATCAC
CACTTGATTATGTAGTAGACTCATTTATAAAAGCAGCTTAGGAACTAATTAACATGGAGGATGAATTACC
TTCTATCCCTTGAGATAAGACATCTTTCAGTTTCATGATTAAGGATTGTTGCTGTTTTATAGTTACTCT
GTTTCATCACAGTGTAATGGTGATGCGTGTCGTAGGTGTGCAGCTATTTGAGGGACTAAGGGATGGAGAT
ATTCTGTCAAATGAATCTCTTCAGTATAACCAGTTTGTGGGAGGGATATGAGACATGTGGATGGCAGTGAG

FIGURE 26.2

63/64

AGATCGTGCCTCTAGATCTTGATGGAGGCTTGGTGAGACACACTTAAATAAGCACGTGGAGGTTAGAATA
GAGGGCAGAGTAAAAGGAAGCTCCATCTGAGCAAGTACACCAAATGATCTCAGCCCTGCAACTTGACCCA
GGTAGGGCCACCACTACGCCTTCACTTGTCAACCAAGCTCCAACCACAGAGAGTTTGACAAGTTTGTGTT
ATGATGTTGGCTTGGCTTGTATTTTAACTTTGGATTTTTAGTGGTTTTGTGCATATAACTGTCTG
AGTTTGGTAGGTAGGATTACTTTGAAAAGGGTTTACTAGTGTGGTCCTCCGGGTAGAATTTAGCTGTAAC
ATGTTGTTAGCCAGCCTGTAGACTGTTAATTACTTAATAATCTCATTGGGAAAATAC TAGTAGTTTTATA
TTTGGATGACATAATTGGAAAAAGCAGATTAGCTGCTACTACTTTTAAAAGACTTAAGGTCGGGATGCCT
TTTTTCCATGTAAGGAAATGAAAAGACCCAAAATCTTCAGGCAAAAGCAAGTTGCAAAATTAGAAACC
ATTGGCTAAAATGTGTTTTGTTGAGTTTCCAAATGGATGAATTTTCATTTGGACATTACATCACTAAAT
TCATTAGATTTTGTCTGCATTGGAAAGATACTCTTCTAGCATATCTTCCCAAAGATATCTAATTTGGAT
TCTGTTTCATGCAAATTTGCATCCCGGAGGTTGAAGTTGGAGTTTGAGGTTGGAAAATATCTTTGAAGGC
AGAATCAGTTGAGTTGTGAGGGTGAAGCCTCACATACTTCTCAACAGACATGATAAAATTCACCTGCATG
AGTTGGCAGGTGGGAGAACCAAACTGGATCACTGGGTAAGACTACTCAGTAAAGCAATGAAGTGCCTTGCT
TAGAGAAGCATCACTATCCCCAAATTTAATTCTGTGGGGAAAAATTAATGAGCCAGTTGTCAAGTCTCTG
ATGGGTCAATTTCAGCACCCCCAAATTTAATTCTGTGGGGAAAAATTAATGAGCCAGTTGTCAAGTCTCTG
TTACATGACTGGCAGACTAAATTTCTTCATCGTTGTTGTTATTTGTTGTTGTTTCTCATTTTCACTTCGC
ACGGCCTTATTCTCATAATTAATTAATTAATTTCTCTTTAGTGTTAGTAGACTCCAACAACAGAAG
TGGCATCTGTGATTATCATAATCAGCATTACCTTGGCAGGAGACTAATCAGATAGGCCGGTCTCAGACAT
TAATCCTACCATCTGATATTTTGGTGAAGGAAAAAGTATTAATTTCTCTTTCCATCCTCCTCCTCAGAAA
TATAGAAGCCCTCTTTACCAAAATCATCACATTTTACTCTGTAATCTACCAGCTAAAAGAAAATTGCATT
GAAGCCCCACAAAGCCAGATTGCAGTTCTTGCCCTTTTTCGCTCTGACATGAGATGTTAAAGAATTATT
CATTGTGCTCACATTGGGTTAGGGGACACTGAACTGCTTTTGTAGATCCATGATCAGTCATCATTCTTCTA
AGAGATTGGAGCTTTGCTGTTTCATTAAGTGTGCAAGTGTAGACTAATGGTGTTAATAAAAAATCATTCAA
AATTTCAAACCTCTTTTGCCAGTGACCTCAATTTGTTGGCTCTGTGATTTGTATCAGACTTTGAGGAGGG
AAGGGGGAAGTGAAGGAAGCCTACGTCCAGGCCCCCTGACAGGATGCTGCAGTAGCAAGCTCAAGCTCGCC
TGCTTGCCAGCAGTTGCTGGTGAGCAGCAGCATGCAGACCAGCTGTGGGAAGCCTCCTGAAGAATGCCCC
AGCTGATGCTTTCAGCTGGGAATAGTTTGTTCCTATTGGGGAACTCATTGTTCTCCAGTCTCTGCAGCAG
GAAGCCAGCTGTCAATTCGGAGGGAATTTAGATGCTTTACCTTTTGGTTTTGTCTGCATCACTCAT
GTGGCTACGAAAGTGTCTCTGAGAATAGAGCCCAATGTGGTGACAATGGGTAGTCAAATGCACCCAGAT
GCTCAAGCCCTGTTGTGTTCTGCAGTGTATGAAATTGGGAGGAAGGAGACCCTGGACAGTAAGCAAA
ATTGGAGACACTCCAACGAGGCTAAGTTAATGCCGTGTTGCCCAGAACAAGATCTAGCTTCTCATTGTT
CAGCCTAGCATGCAACCAAGTGGTGTGCTGGTAAATGTTTAAACAACAGCTCGCTGAGAATAGAAAGCAC
CTGTTTTGCACCAATTTGCCAATTTCCATGGCATAAATACTACCCTTTAGATGATTTAAGCTACCAACT
GTGATGTCACTGAACACATGGTTGGAAAGAGATGCACGCAGTTGGCTCTTGCAAGCCTGGGCAAAAATGC
TTCAACACGCCACTGGATGCAGCCAGTCAGAGGGTTTCAATTTAATATATGTGTTTCATGTGGACACACAC
AGACACACACACACAACTCACCTTACACACACACTTCGATGACTAAAAACAATTACATAGTTTTTAAGAT
ATGAATCAATGTGTGAATGTAGAAAGCTTATGATAAGGCCCTAGAGGTATGGGTTGCCCTGGAAGCCTAG
GTTTTAAGCAGGAGAATAGCTGAGAAGAATGAAGCCCTCCTGAGCTGAAAGGAGAGATGGATCAATGGAG
ATGGTTCCATCATCTCCTTCCATATCTCACAGGTAAAATGGGCACTCAGAAAACCCTCACGATTGATTTT
TTAAAAAGATAAGTGAGTGTTTTTTATTTTATTATTATTGTGCATCATTATTTTGATTTACAAATGCTATT
TGTAACTTTTTACATGTAAGTAAAGTATTTACGGGAACCTCTATGGAGAATAGCACAATCCAGAATT
TACTGTGTTTTTCTTTTATGTGACGTGGAAACTCAGTAATTTCTCCACCTTCACATTGTTGTTTCAATA
ATTTTACTTTTAGTTATTAGGGAATCTAAGTTTTTGTAAACATTTGTTTTTGTAAAGTATCTACTTA
CTGTTTTAGCTCTGAACCTCAAACCAGAATATCTCTGTATCAATTGCATGACTATTTCAGAAACAATAATCC
AAACCAAAATAATTTCTTTTCCACCCAGTACGAAGAAAATAAGCTCAGTAACAAGAAGGCATAAACTAA
AGTATATAATGAGGCTTTCATTAAATACACACACACACACTCACACACACACATACACTTTTTTAA
TTTTTAAATTAGGCCTCCACACATAAATCATTTTGAAAGTAGAATAGAAAATCTCAAAGAATTCATTCTC
CTGGTCTGTGCATCTTCTGCAGTTAATAAGAGGTTTGTATCTGGAAAGATGGAAGAACTTGTTCTAAAA
TCTTATTTTTTCAAAAAAAATTTCCATTTCTCTCTGGGCCTGTATCCATGGTTGAATGTTAGCCCTGGA
GGAGATCCATGCTTACTCGCTCTTTCTGGCCCTTCTGTCTTTGCCTCTGCAATTTCTTTTGTAGCTGG
CACGATAGCAGGGACTGGGGGTCTATCCTTTTCATGGTATTGCTACAATATTTGTCTTACTGGAAGTGG
TAACATCCGGGTCTGATTTAATTGGCATTACACTTACACAGGGACTCTGAGCACCCCCGTCACCACACCA

FIGURE 26.3

64/64

GACAGTGGACCAGTTTTTCACAGCTACAAAGAGCTAGAAATGTGTTTAAACATCATCCAGTGCATCCCCTAA
TTCAAAACCATCCTCACTAATCAATCATATTACCCATAAATATTACAAATGAGATTGATTCCATCTCAA
GACAATTTGTCAAATACTTAATTTTCTTCCTGGATGATTCTACTTACTGGATATTTTAGAAAGAGAAATG
TCTGAGATAAAATCCCTCACATTTACTCAATATAACAAATTACTGTTTCTACTCCTATTCTGAGTAGTGC
TTCTGAAGATTGTTTGGCTGTAGTGTGTCTTTGATAAAATGAATGTCAGTAGTGAGCCTTTTAGAGATAC
CATGCTCAGACATCCTCTTTGGGATCAGAAGATACCTAAATTTCTCCCCTTTTGCCCACTTGGTTAGATG
AGTGATATATTCTTTGGATCCTGCAAAGAAGAGATTGGTTTCTTTTCTTTTCTGGTGGTGGTAGTGGTTG
TATCTGTGGCTGTGATGGTTGTTGTTACTTGTCTCTCTCTCTCTCTGGCTCTGGCTTTTGCTTTCTCTGCT
AGTGTCTTTTCTCTTTCCAAACAAATAGTTAAATTAAACGTGAGCTTCTGAATTGTACTTGTTCATACTT
TCAAAACATAACAGATTAAATAAAATAGATGTGTCTGATTTAAACATGCCCCCTGGAAAGGCATGCTG
TATTATGAAATCGTGATAATACTGCATTATTACATGGCAGTATAAATATTAGTCTGTTGAATTCATT
TGTCCAATTGTATAACTTTGTGGAGCAGTGTTTTGACCTTTGATACATAATTCTGGAGCAAGTGGAGTGG
TTGCAGGCAGATGAGACAGTGTATATCAGGATTTTCAATCAACTTTAGTTGGAGGCCTGGCAATTACA
AACATCTTCAGATGTTTCTGTAACCATTATAAATATGAAAAAACCTCTTCAAAAAATTTCCCATAGTAC
TTCAGTCAAGACTTTTTAGGTTTATCTTTTTTTTTTTCATTTCTCCTTTTCCCTTTTCCATTATTTTTCGAT
GGGGGGGTTGTTATCATTGACTGAAGAAATATTTTGATTGCAATGGTCTCTCTCTCTCTCCCCCTCTCTC
TCTCTCTCCTCTATTCTTTCCCTCCTTCCCTCTGTCCATCACCCCTCATTAAAAATATTGAAATCTGGAGTC
TTTGATAAATCTGCATTAGACCAGGCTATATGCTAGGAATGAAATCTGGGCAAATATCGATGGGTTTTCA
AAGAATGCTCCATGTTTATTGGGCCCTTTACACCCACAGTGATAAATGAAAAGGATAGAGGTAGTTTT
TTCAAAAGAGCACTTTAATAATATCCTCTGAGACCTAATGCAGTTTAAACAAATGACTCCACCTATTTTTTC
CAGTAGGTAAATTGACTGAGACTTGCAAAATACCCCTGAGAGTTGTGAGGGGTGTCTTCTGCCTGGTCTA
TAGCGTGTGTGTTTGTCTTTGTATCTAACAGGCACATTACGTCCTCGTGTACTCATATGAAGTATTTCTTA
ACATTCCCATTAGCCTGTATATAAGAATCAGAAAGATAATCCCAACATGTTGTAAATGAAGATGTGACTC
TATAACCTTTCTCTTCTTCTGGAAGAAAAAGGACATTTTCATGCATATTTTAAACAGAAATTTTGTATA
TTTAAGTGTCATAGAAAAATATTTATTGAGTAACCTGGGACACAAATGGGAATTTAATTGTTCATCATATGCT
TTGTGTGTGGGGATGCTTACCAACACCATGTGCTGCTGGACCAATTGTGGCAAGCCATAACTGCACAAAGAGT
ACACATCGTCAGTGTGTGTGTGTGTGTGTGTGTGCGCGCACGCACGTGCGTGTGTGTGTCCCTGCATGTG
CAACATGTCTAGCTTGCTGTCTTTCATGGGATTTTAGCTTTCCCTTCTTGAAAAACATTATTTTACAGTT
CCAGGAGGCCCTGGTTACATTACTATATGAAGGCAGTGATTTGAAATGAAAATTCCTTTCTCTTGAAG
CTTTGGTCATAATATCATGGTTCAATTAAACGGATTCCACCGGACTTTGTGATGAAAAAGGCTCTGTTAA
AATCCAATTGAGTTTCCAAGAGGAAATTGTAGTAGGTCAAGATGCATGAGAGGGAAAGATGGAGGCCACCT
CAGCTGGAGAACATGAGCTGAGTTGAGCCCTCAGTGTGAAAGTTGACTTGCTCCAAGCTGCAGTCTAAAA
CCCTGGGGCCCGTGCTTGGCTATGCTCCCTCCCAAGTAAGTAGAGGAGCAGAACCATCAGGAACAGCCT
GCCTGGCTCCTATGAAGAAAACCTCCTGACGTCTGTCCCCAAAGGAAGACCCTTTCCCCAAGGGCACCC
CAGGTGGCCATTAAATTTGTGATGATCATTCAGAAAGTGCCCCCTTGGCTTTATGAGAATCCAATTAGTCT
TCTGAACCACCTTTTCTTGGGTGCAGATTTCCAACATTTCATGCTCATTGCAGATCCACCAACTGTCACCTG
TTCTTAAACAAGCATGCTCGTCTTGTGAGAAATTCAGTAAGTTCCAATTTCTGTACAGACCAGGGTAAAC
TGTTCTAAATCAATCAATTAATGAAATGTTATCTGGTTTTTAAAGCTGGTTTTCATGTGCTTTATGTGT
ATAAACTATATCTGCCTGTGTGGCTTTGCATTTCAAATGTGTGGCGCACAAAGCGTTTTGTTGGTGCTTT
GTTCTCAGTACAGTAACCTCTGTGTACAAACATTTTAAATGTGGTTTTGTTGTTTTCCAACAAGATGTCTCT
GTAAAAATGATATTGGCTGAGCTGGTGGCTTGGTTTTCTCTCATAGAGGCATTAACTATACTGCCAATGCA
TTGAATTATTTAAAAATGCAAAATAAAATTTTTATGAAAATCTCA

FIGURE 26.4

LISTE DE SEQUENCES

<110> INSERM
ALLIEL, Patrick
PERIN, Jean-Pierre
RIEGER, Francois

<120> FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES
PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX
ENDOGENES HUMAINS, LEURS SEQUENCES FLANQUANTES ET LEURS
APPLICATIONS.

<130> 598EXT21

<140>
<141>

<150> 9807920
<151> 1998-06-23

<160> 122

<170> PatentIn Ver. 2.1

<210> 1
<211> 2599
<212> ADN
<213> Homo sapiens

<400> 1

atcccctgcc	ttaatcgcca	agctccttca	ggagaacaaa	gaacaggcca	ttaccctgga	60
gaagactggc	aactgatttt	accacaagc	ccaaacctca	gggatttcag	tatctactag	120
tctgggtaga	tactttcacg	ggttgggcag	aggccttccc	ctgtaggaca	gaaaaggccc	180
aagaggtaat	aaaggcacta	gttcatgaaa	taattccag	attcggactt	ccccaggcgt	240
tacagagtga	caatagccct	gctttccagg	ccacagtaac	ccaggaggta	tcccaggcgt	300
taggtatacg	atatcactta	cactgcgct	gaaggccaca	gtcctcaggg	aaggtcgaga	360
aaatgaatga	aacactcaaa	ggacatctaa	aaaagcaaac	ccaggaaacc	cacctcacat	420
ggcctgctct	gttgccctata	gccttaaaaa	gaatctgcaa	ctttcccca	aaagcaggac	480
ttagcccata	cgaaatgctg	tatggaaggc	ccttcataac	caatgacctt	gtgcttgacc	540
caagacagcc	aacttagttg	cagacatcac	ctccttagcc	aaatatcaac	aagttcttaa	600
aacattacaa	ggaacctatc	cctgagaaga	gggaaaagaa	ctattccacc	cttgtgacat	660
ggtattagtc	aagtcccttc	cctctaattc	cccatcccta	gatacatcct	gggaaggacc	720
ctaccagtc	attttatcta	ccccaaactgc	ggttaaagtgc	gctggagtgg	agtcttggt	780
acatcacact	tgagtcaaat	cctggatact	gccaaaggaa	cctgaaaatc	caggagacaa	840
cgctagctat	tctgtgaac	ctctagagga	tttgcgctg	ctcttcaaac	aacaaccagg	900
aggaaagtaa	ctaaaatcat	aaatccccat	ggccctccct	tatcatattt	ttctctttac	960
tggtctttta	ccctctttca	ctctcactgc	acccctcca	tgccgctgta	tgaccagtag	1020
ctccccttac	caagagtttc	tatggagaat	gcagcgctcc	ggaaatattg	atgccccatc	1080
gtataggagt	ctttctaagg	gaacccccac	cttcactgcc	cacacccata	tgccccgcaa	1140
ctgctatcac	tctgccactc	tttgcatgca	tgcaaatact	cattattgga	caggaaaaat	1200
gattaatcct	agttgtcctg	gaggacttgg	agtcactgtc	tggttgactt	acttcaccca	1260
aactggtagt	tctgatgggg	gtggagtcca	agatcaggca	agagaaaaac	atgtaaaaga	1320
agtaatctcc	caactcacc	gggtacatgg	cacctctagc	ccctacaaag	gactagatct	1380
ctcaaaacta	catgaaaccc	tccgtaccca	tactcgctg	gtaagcctat	ttaataccac	1440
cctcactggg	ctccatgagg	tctcggccca	aaaccctact	aactgttgga	tatgcctccc	1500
cctgaacttc	aggccatctg	tttcaatccc	tgtacctgaa	caatggaaca	acttcagcac	1560
agaaataaac	accactccg	ttttagtagg	acctctgtgt	tccaatctgg	aaataaccca	1620
tacctcaaac	ctcacctgtg	taaaatttag	caatactaca	tacacaacca	actcccaatg	1680

```

catcagggtgg gtaactcctc ccacacaaat agtctgccta cctcaggaa tattttttgt 1740
ctgtgggtacc tcagcctatc gttgtttgaa tggctcttca gaatctatgt gcttcctctc 1800
attcttagtg cccctatga ccatctacac tgaacaagat ttatacagtt atgtcatatc 1860
taagccccgc aacaaaagag taccattct tctttttgtt ataggagcag gagtgttagg 1920
tgcactagggt actggcattg gcggtatcac aacctctact cagttctact acaaaactatc 1980
tcaagaacta aatggggaca tggaacgggt cgccgactcc ctggtcacct tgcaagatca 2040
acttaactcc ctagcagcag tagtccttca aaatcgaaga gcttttagact tgctaaccgc 2100
tgaaagagggt ggaacctgtt tatttttagg ggaagaatgc tgttattatg ttaatcaatc 2160
cggaatcgtc actgagaaag ttaaagaaat tcgagatcga atacaacgta gagcagagga 2220
gcttcgaaac actggaccct ggggcctcct cagccaatgg atgcccgtga ttctcccctt 2280
cttaggacct ctagcagcta taatattgct actcctcttt ggacctgta tctttaacct 2340
ccttgtaaac tttgtctctt ccagaatcga agctgtaaaa ctacaaatgg agcccaagat 2400
gcagtcceaag actaagatct accgcagacc cctggaccgg cctgctagcc cacgatctga 2460
tgtaaatgac atcaaaggca cccctcctga ggaaatctca gctgcacaa ctctactacg 2520
ccccatttca gcaggaagca gttagagcgg tctcggccaa cctccccaac agcacttagg 2580
tttctctgtt gagatgggg 2599

```

<210> 2

<211> 1326

<212> ADN

<213> Homo sapiens

<400> 2

```

gccgcctggc actcctgagg gaagtataaa ttataacacc atcttacagc tagacctctt 60
ttgtagaaaa ggcaaatgga gtgaagtgcc ataagtacaa actttctttt cattaagaga 120
caactcacia ttatgtaaaa agtgtgattt atgccctaca ggaagccttc agagtctacc 180
tccctatccc agcatccccg actccttccc caactaataa ggacccccct tcaaccctaaa 240
tggtccaaaa ggagatagac aaaagggtaa acagtgaacc aaagagtgcc aatattcccc 300
aattatgacc cctccaagca gtgggaggaa gagaattcgg ccagccaga gtgcatgtgc 360
ctttttctct cccagactta aagcaataaa aaacagactt aggtaaattc tcagataacc 420
ctgatggcta tattgatgtt ttacaagggt taggacaatt ctttgatctg acatggagag 480
atataatgtc actgctaaat cagacactaa ccccaatga gagaagtgcc accataactg 540
cagcctgaga gtttggcgat ctctggtatc tcagtcaggt caatgatagg atgacaacag 600
aggaaagaga atgattcccc acaggccagc aggcagttcc cagtctagac cctcattggg 660
acacagaatc agaacatgga gattggtgct gcagacattt gctaacttgt gtgctagaag 720
gactaaggaa aactaggaag aagtctatga attactcaat gatgtccacc ataacacagg 780
gaaggggaaga aaatcctact gcctttctgg agagactaag ggaggcattg aggaagcgtg 840
cctctctgtc actgactct tctgaaggcc aactaatctt aaagcgtaaag tttatcactc 900
agtcagctgc agacattaga aaaaaacttc aaaagtctgc cgtaggcccc gaggcaaaact 960
tagaaacctt attgaacttg gcaacctcgg ttttttataa tagagatcag gaggagcagg 1020
cggaacagga caaacgggat taaaaaaaaa gccaccgctt tagtcatgac cctcaggcaa 1080
gtggactttg gaggtctctg aaaagggaaa agctgggcaa attgaatgcc taatagggtc 1140
tgcttccagt gcggtctaca aggacacttt aaaaaagatt gtccaagtag aagtaagccg 1200
ccccctcgtc catgccctt atttcaagg aatcactgga aggccactg cccagggga 1260
caaaggtcct ctgagtcaga agccactaac cagatgatcc agcagcagga ctgagggtgc 1320
ctgggg 1326

```

<210> 3

<211> 10499

<212> ADN

<213> Homo sapiens

<400> 3

```

ccctggggcg ggcttccttt ctgggatgag ggcaaacgc ctggagatac agcaattatc 60
ttgcaactga gagacaggac tagctggatt tcctaggccg actaagaatc cctaagccta 120
gctgggaagg tgaccacgtc cacctttaa caccgggctt gcaacttagc tcacacctga 180

```

ccaatcagag	agctcactaa	aatgctaatt	aggcaaagac	aggaggtaaa	gaaatagcca	240
atcatctatt	gcctgagagc	acagcaggag	ggacaacaat	cgggatataa	acccaggcat	300
tcgagctggc	aacagcagcc	cccccttggg	tccttccct	ttgtatggga	gctgttttca	360
tgctatttca	ctctattaaa	tcttgcaact	gcactcttct	ggcccatgtt	tcttacggct	420
cgagctgagc	ttttgctcac	cgctccaccac	tgctgtttgc	caccaccgca	gacctgccc	480
tgactcccat	ccctctggat	cctgcagggt	gtccgctgtg	ctcctgatcc	agcgaggcgc	540
ccattgccgc	tcccaattgg	gctaaaggct	tgccattgtt	cctgcacggc	taagtgcctg	600
ggtttgttct	aattgagctg	aacactagtc	actgggttcc	atggttctct	tctgtgaccc	660
acggcttcta	atagaactat	aacacttacc	acatggccca	agattccatt	ccttggaatc	720
cgtgagggca	agaactccag	gtcagagaa	acgaggcttg	ccaccatctt	ggaagcggcc	780
tgctaccatc	ttggaagtgg	ttcaccacca	tcttgggagc	tctgtgagca	aggaccccc	840
ggtaacattt	tggcaaccac	gaacggacat	ccaaagtgg	gagtaatatt	ggaccacttt	900
cacttgctat	tctgtcctat	ccttccttag	aattggagga	aaataccggg	cacttgctcg	960
ccagttaaaa	acgattagtg	tggccaccgg	acttaagact	cagggtgtgag	gctatctggg	1020
gaagggcttt	ctaacaaccc	ccaacccttc	tgggttgggg	acttggtttg	cctcaagcca	1080
gcttcacatt	tacgttttct	tggggaagcc	gagggccgac	tagaggcaga	aagctgtctg	1140
cctgaactcc	cggcagtagc	cggttgagat	catgggtgtg	ccagaagtct	caacagtcgc	1200
ccatgcatgc	acccctatct	ttccttctga	cccatacctc	ctgggtccca	accacaactt	1260
tcttcaaagt	gtagccccaa	aattctcctt	acctctgaat	atacttctct	tgatccctgc	1320
ctcctaggta	ctattggttc	agacttccat	ttcctctagc	aagttgtatc	tccaaaaggga	1380
tctaaggaag	ctctgcgctg	cgctccttag	cacctaggct	ataaccagag	gagtccttatc	1440
cctggtgtcc	ctcccaattt	aggcatacag	ctcttgacat	gggcagttat	gtaggacca	1500
ctccccacca	cccttgccag	ggccccaagt	ttgtaaatgg	ctgagggaaa	agagagacag	1560
aggagagaga	gagaaatgga	ggagaaagag	agagagacag	agaggagaga	gagacagtga	1620
gagagacaga	agagagagag	agacaaagag	gagagagaga	gagtcaaaga	gagaaagaaa	1680
gagaaagaaa	tagtaaaaaa	cagtgtgccc	tattccttta	aaagccaggg	taaattttaaa	1740
acctgtactt	gataattgaa	ggtcttctct	gtgaccctat	agcactccaa	tccactttgt	1800
ggtcagtgta	aataagagca	taggccgaaa	gcactgaggc	cattgacaac	ccgtagcttc	1860
cctatcaaaa	atccttaacc	cagtaaccgc	cagatggacc	aaatgcattc	agtcggtagc	1920
gcaactgctt	tgctaaaagt	agaaaagtaa	cttttagagg	aaacctcatt	gtgagcacac	1980
ctcacctgtt	cagaattatt	ctaataaaaa	aagcaaaaag	gtagcttact	aactcaaaaa	2040
tcttaaagta	tggggctatt	ctgttagaaa	aaggtaatgt	aactccaacc	actgataatt	2100
cccttaaccc	agcagatttc	ctaaccggat	ttaaacttta	attaccatac	aaaggtccga	2160
ccagacttag	gcggaactcc	cttcaggaca	ggacgataga	tggttcctcc	tggttctctg	2220
aggaaaaaaa	ccacaatggg	tattcagtaa	ttgatacggg	gactcttgtg	gaagcagagt	2280
tagaaaaatt	gcctaataac	tggtctcctc	aaacgtgtga	gctgtttgca	ctcagccaag	2340
ccttaaagta	cttacagaat	caaaagacta	tctcaatcct	gattcaaaaag	gttagctaca	2400
ccctctctgt	aatgcatttg	cataagaact	tgtttatggg	aatgcattct	gatggggcag	2460
ctgggttgtt	ataaaaatag	aaccagcccc	agctctagga	ctcacccttg	agcgcaaatg	2520
caatgttggg	catgctggta	aaggaccact	agaatccagc	agcccagacc	cttttctttg	2580
tgggtcaagaa	aggcgggaaa	aggggtgcag	gactgctaca	tcggtaagca	taactaatcc	2640
gataaacaga	ggtccatggg	tggttacgca	ccctggaaag	gaactcacc	ctgagcaca	2700
aggcaatgtt	gggcacgctg	gtaaaggacc	actagaatcc	agcagcctgg	acccctttct	2760
ttgtgggtcaa	gagaggcagg	aaaacagggt	caggactgca	acatcagtga	gcataactaa	2820
ttcgataagc	agaggtccat	gggtggtgat	gcaccctgga	agaataagc	attaggacca	2880
tagaggacac	tccaggacta	aagctcatcg	gaaaatgact	agggttgctg	gcatccctat	2940
gttctttttt	cagatgggaa	acgttccccg	caagacaaaa	acgcccctaa	gacgtattct	3000
ggagaatttg	gaccaatttg	accctcagac	actaagaaag	aaacgactta	tattcttctg	3060
cagtgccgccc	tggcactcct	gagggaaagta	taaattataa	caccatctta	cagctagacc	3120
tctttttag	aaaaggcaaa	tggagtgaag	tgccataagt	acaaactttc	ttttcattaa	3180
gagacaactc	acaattatgt	aaaaagtgtg	attttatgcc	tacaggaagc	cttcagagtc	3240
tacctcccta	tcccagcatc	cccgaactcct	tcccacta	ataaggaccc	cccttcaacc	3300
caaattggtcc	aaaaggagat	agacaaaagg	gtaaacagtg	aaccaaagag	tgccaatatt	3360
cccccaattat	gacccctcca	agcagtggga	ggaagagaa	tcggcccagc	cagagtgcac	3420
gtgccttttt	ctctcccaga	cttaaagcaa	ataaaaacag	acttaggtaa	attctcagat	3480
aaccctgatg	gctatattga	tgttttacaa	gggttaggac	aattctttga	tctgacatgg	3540
agagatataa	tgctactgct	aaatcagaca	ctaaccctaa	atgagagaag	tgccaccata	3600
actgcagcct	gagagtttgg	cgatctctgg	tatctcagtc	aggtcaatga	taggatgaca	3660

acagaggaaa	gagaatgatt	ccccacaggc	cagcaggcag	ttcccagtc	agaccctcat	3720
tgggacacag	aatcagaaca	tggagattgg	tgctgcagac	atttgctaac	ttgtgtgcta	3750
gaaggactaa	ggaaaactag	gaagaagtct	atgaattact	caatgatgtc	caccataaca	3840
caggggaagg	aagaaaatcc	tactgccttt	ctggagagac	taagggaggc	attgaggaa	3900
cgtgcctctc	tgtaacctga	ctcttctgaa	ggccaactaa	tcttaaagcg	taagtttatc	3960
actcagtcag	ctgcagacat	tagaaaaaaa	cttcaaaagt	ctgccgtagg	ccgggagcaa	4020
aacttagaaa	ccctattgaa	cttggcaacc	tccgtttttt	ataatagaga	tcaggaggag	4080
caggcggaac	aggacaaacg	ggattaaaaa	aaaggccacc	gcttttagtca	tgaccctcag	4140
gcaagtggac	tttggaggct	ctggaaaagg	gaaaagctgg	gcaaattgaa	tgccataatg	4200
ggcttgcttc	cagtgcggtc	tacaaggaca	ctttaaaaaa	gattgtccaa	gtagaagtaa	4260
gccgccccct	cgtccatgcc	ccttattttca	agggaatcac	tggaaggccc	actgccccag	4320
gggacaaagg	tctctgagt	cagaagccac	taaccagatg	atccagcagc	aggactgagg	4380
gtgcctgggg	caagcgccat	cccatgccat	caccctcaca	gagccctggg	tatgcttgac	4440
cattgagggc	caggagggtt	tctcctggag	actggtgcgg	tcttcttagt	cttactcttc	4500
tgtcccggac	aactgtcctc	cagatctgtc	actatctgag	ggggtcctaa	gacgggcagt	4560
cactagatac	ttctcccagc	cactaagtta	tgactgggga	gctttattct	tttcacatgc	4620
ttttctaatt	atgcttgaaa	gccccactac	cttgttaggg	agagacattc	tagcaaaagc	4680
aggggccatt	atacacctga	acataggaga	aggaacaccc	gtttgttgtc	ccctgcttga	4740
ggaaggaatt	aatcctgaag	tctgggcaac	agaaggacaa	tatggacgag	caaagaatgc	4800
ccgtcctggt	caagttaaac	taaaggattc	cacctccttt	ccctaccaa	ggcagtaccc	4860
cctcagaccc	aaggcccaac	aaggactcca	aaagattgtt	aaggacctaa	aagcccaagg	4920
cctagtaaaa	ccatgcagta	acccctgcag	tactccaatt	ttaggagtag	agaaacccaa	4980
cagacagtgg	aggttagtgc	aagatctcag	gattatcaat	gaggctgttg	ttcctctata	5040
gccagctgta	cctagccctt	atactctgct	ttcccaaata	ccagagggaag	cagagtgggt	5100
tacagtccct	gaccttcagg	atgccttctt	ctgcatccct	gtacatccct	actctcaatt	5160
cctgtttgcc	tttgaagata	cttcaaacc	aacatctcaa	ctcacctgga	ctattttacc	5220
ctaagggttc	agggatagtc	cccatctatt	tgccaggga	ttagcccaag	acttgagcca	5280
atcctcatac	ctggacactt	gtccttcggg	agggtgtaga	tttacttttg	gccgcccttc	5340
cagaaacctt	gtgccatcaa	gccacccaag	cgtctttcaa	tttctctcgt	acctgtggct	5400
acatgggttc	caaaccaaag	gctcaactct	gctcacagca	ggttacttag	ggctaaaatt	5460
atccaaaggc	accaggggcc	tcagttagga	acacatccag	cctatactgg	cttatccctca	5520
tcccaaaacc	ctaaagcaac	taaggggatt	cctttggcgt	atagggtttct	gccgaaaatg	5580
gattcccagg	tatggcgaaa	tagccaggtc	attaaatata	ctaattaaag	aaactcagaa	5640
agccaatacc	catttagtaa	gatggacaac	tgaagtagaa	gtggctttcc	aggccctaac	5700
ccaagcccca	gtgttaagtt	tgccaacagg	gcaagacttt	tcttcataat	tcacagaaaa	5760
aacaggaata	gctctaggag	tccttacaca	gatccgaggg	atgagcttgc	aacctgtggc	5820
atacctgact	aaggaaattg	atgtagtggc	aaagggttga	cctcattgtt	tacgggtagt	5880
ggtggcagta	gcagtcttag	tatctgaagc	agttaaaata	atacagggaa	gagatcttac	5940
tgtgtggaca	tctcatgatg	tgaatggcat	actcactgct	aaaggagact	tgtggctgtc	6000
agacaactgt	ttacttaaat	gtcaggctct	attacttgaa	gggccagtgc	tgcgactgtg	6060
cacttgtgca	actcttaacc	cagccacatt	tcttcagac	aatgaagaaa	agataaaaaca	6120
taactgtcaa	caagtaattt	ctcaaacct	tgccactcga	ggggaccttt	tagaggttcc	6180
tttgactgat	cccgacctca	acttgtatac	tgatggaagt	tcctttgtag	aaaaaggact	6240
tcgaaaagt	gggtatgcag	tggtcagtga	taatggaata	cttgaaagta	atccctcac	6300
tccaggaact	agtgtcagc	tagcagaact	aatagccctc	acttgggcac	tagaattagg	6360
agaagaaaaa	agggcaaata	tatatacaga	ctctaaatat	gcttacctag	tcctccatgc	6420
ccatgcagca	atatggaaag	aaagggaatt	cctaacttct	gagagaacac	ctatcaaaca	6480
tcaggaagcc	attaggaat	tattattggc	tgtacagaaa	cctaaagagg	tggcagtctt	6540
acactgccgg	ggtcatcaga	aaggaaaagga	aagggaaata	gaagagaact	gccaagcaga	6600
tattgaagcc	aaaagagctg	caaggcagga	ccctccatta	gaaatgctta	taaaacaacc	6660
cctagtatag	ggtaatcccc	tccgggaaac	caagccccag	tactcagcag	gagaaacaga	6720
atggggaaac	tcacgaggac	agttttctcc	cctcgggacg	gctagccact	gaagaaggga	6780
aaatactttt	gcctgcaact	atccaatgga	aattacttaa	aacccttcac	caaacccttc	6840
acttaggcac	cgatagcacc	catcagatgg	ccaaatcatt	atttactgga	ccaggccttt	6900
tcaaaactat	caagcagata	gtcagggcct	gtgaagtgtg	ccagagaaat	aatccctgc	6960
cttatcgcca	agctccttca	ggagaacaaa	gaacaggcca	ttaccctgga	gaagactggc	7020
aactgatttt	accacaagc	ccaaacctca	gggattttcag	tatctactag	tctgggtaga	7080
tacttttcacg	ggttggggcag	aggccttccc	ctgtaggaca	gaaaaggccc	aagaggtaat	7140

aaaggcacta	gttcatgaaa	taattcccag	attcggactt	ccccgaggct	tacagagtga	7200
caatagccct	gctttccagg	ccacagtaac	ccagggaagta	tcccaggcgt	taggtatacg	7260
atatcactta	cactgcgcct	gaaggccaca	gtcctcaggg	aaggtcgaga	aaatgaatga	7320
aacactcaaa	ggacatctaa	aaaagcaaac	ccaggaaacc	cacctcacat	ggcctgctct	7380
gttgccata	gccttaaaaa	gaatctgcaa	ctttcccca	aaagcaggac	ttagcccata	7440
cgaaatgctg	tatggaaggc	ccttcataac	caatgacctt	gtgcttgacc	caagacagcc	7500
aacttagttg	cagacatcac	ctccttagcc	aaatatcaac	aagttcttaa	aacattacaa	7560
ggaacctatc	cctgagaaga	gggaaaagaa	ctattccacc	cttgtgacat	ggtattagtc	7620
aagtcccttc	cctctaattc	cccatcccta	gatacatcct	gggaaggacc	ctaccagtc	7680
attttatcta	ccccaaactgc	ggttaaagtg	gctggagtgg	agtcttgat	acatcacact	7740
tgagtcaaat	cctggatact	gccaaaggaa	cctgaaaatc	caggagacaa	cgctagctat	7800
tcctgtgaac	ctctagagga	tttgcgctg	ctcttcaaac	aacaaccagg	aggaaagtaa	7860
ctaaaatcat	aaatcccat	ggccctccct	tatcatat	ttctctttac	tgttctttta	7920
ccctctttca	ctctcactgc	acccctccca	tgcgcgtgta	tgaccagtag	ctccccctac	7980
caagagtttc	tatggagaat	gcagcgtccc	ggaaatattg	atgcccctac	gtataggatg	8040
ctttctaagg	gaacccccac	cttcaactgcc	cacacccata	tgccccgcaa	ctgctatcac	8100
tctgccactc	tttgcatgca	tgcaaatact	cattattgga	caggaaaaat	gattaatcct	8160
agttgtcctg	gaggacttgg	agtcaactgtc	tggtggactt	acttcaccca	aactgggatg	8220
tctgatgggg	gtggagtcca	agatcaggca	agagaaaaac	atgtaaaaga	agtaatctcc	8280
caactcacc	gggtacatgg	cacctctagc	ccctacaaag	gactagatct	ctcaaaacta	8340
catgaaaccc	tccgtaccca	tactcgctg	gtaagcctat	ttaataccac	cctcactggg	8400
ctccatgagg	tctcggccca	aaaccctact	aactgttggga	tatgcctccc	cctgaacttc	8460
aggccatatg	tttcaatccc	tgtacctgaa	caatggaaca	acttcagcac	agaaataaac	8520
accacttccg	tttttagtagg	acctcttggt	tccaatctgg	aaataaccca	tacctcaaac	8580
ctcacctgtg	taaaatttag	caatactaca	tacacaacca	actcccaatg	catcagggtg	8640
gtaactcctc	ccacacaaat	agtctgccta	ccctcaggaa	tattttttgt	ctgtggtacc	8700
tcagcctatc	gttgtttgaa	tggctcttca	gaatctatgt	gcttcctctc	attcttagtg	8760
ccccctatga	ccatctacac	tgaacaagat	ttatacagtt	atgtcatatc	taagccccgc	8820
aacaaaagag	taccattctc	tccttttggt	ataggagcag	gagtgcctag	tgactagggt	8880
actggcattg	gcggtatcac	aacctctact	cagttctact	acaaactatc	tcaagaacta	8940
aatggggaca	tggaacgggt	cgccgactcc	ctggtcacct	tgcaagatca	acttaactcc	9000
ctagcagcag	tagtctctca	aaatcgaaga	gctttagact	tgctaaccgc	tgaaagaggg	9060
ggaacctgtt	tatttttagg	ggaagaatgc	tgttattatg	ttaatcaatc	cggaatcgct	9120
actgagaaag	ttaaagaaat	tcgagatcga	atacaacgta	gagcagagga	gcttcgaaac	9180
actggaccct	ggggcctcct	cagccaatgg	atgccctgga	ttctcccctt	cttaggacct	9240
ctagcagcta	taatattgct	actcctcttt	ggaccctgta	tctttaacct	ccttggtaac	9300
tttgtctctt	ccagaatcga	agctgtaaaa	ctacaaatgg	agcccaagat	gcagtccaag	9360
actaagatct	accgcagacc	cctggaccgg	cctgctagcc	cacgatctga	tgtaaatgac	9420
atcaaaggca	ccctcctga	ggaaatctca	gctgcacaac	ctctactacg	cccccaattca	9480
gcaggaagca	gttagagcgg	tctcggccaa	cctccccaac	agcacttagg	ttttcctggt	9540
gagatggggg	actgagagac	aggactagct	ggatttccta	ggctgactaa	gaatccctaa	9600
gcctagctgg	gaagggtgacc	acatccacct	ttaaacacgg	ggcttgcaac	ttagctcaca	9660
cctgaccaat	cagagagctc	actaaaatgc	taattaggca	aagacaggag	gtaaagaaat	9720
agccaatcat	ctattgcctg	agagcacagc	aggagggaca	atgatcggga	tataaaccca	9780
agtcttcgag	ccggcaacgg	caacccccct	tgggtcccct	ccctttgtat	gggagctctg	9840
ttttcatgct	atttcaactc	attaaatctt	gcaactgcac	tcttctggtc	catgtttctt	9900
acggcttgag	ctgagctttc	gctcgccatc	caccactgct	gtttgccgcc	accgcagacc	9960
cgccgctgac	tcccatccct	ctggatcatg	cagggtgtcc	gctgtgctcc	tgatccagcg	10020
aggcacccat	tgccgctccc	aatcgggcta	aaggcttgcc	attgttctctg	catggctaag	10080
tgccctgggtt	catcctaatt	gagctgaaca	ctagtcaactg	ggttccatgg	ttctcttctg	10140
tgaccacacag	cttctaatag	agctataaca	ctcaccgcat	ggcccaagggt	tccattcctt	10200
gaatccataa	ggccaagaac	cccaggctcag	agaacacgag	gcttgccacc	atcttgggag	10260
ctctgtgagc	aaggaccccc	aagtaacaca	accatgaggg	tgcaaatgca	tgggccacta	10320
atggtagagc	aagaaaacag	aagggccctg	gttcctcgaa	ggcatcagtg	agctgaaatg	10380
cctgccctgg	atgtcctatt	cctaggtggt	tttctgctg	aagcagatta	aaccctttgt	10440
tcactttctcc	aagtagggct	tctattacag	cccaaatcaa	tccccacccc	agatgacat	10499

<210> 4
 <211> 2784
 <212> ADN
 <213> Homo sapiens

<400> 4
 ctccttcagg agaacaaaga acaggccact acccaagaga agactggcaa ctagatttta 60
 cccatatgcc caaatctcag ggattttcagt atctactagt ttgggtagat acttttactg 120
 gttgggcaga ggccttcccc tgtaggacag aaaaggccca agaggtaata aacgttcatg 180
 aaataattcc cagattcgga cttccccaag gcttacagag tgacaatggc cctgctttca 240
 aggctacagt aacccaagga gtatcccagg tgttagggtat acaatatcac tcacactgcg 300
 cctggaggcc acagtccctca ggaaagggtg agaaaatgaa caaaacactc aaatgacatc 360
 taaaaaagct aatccaggaa acccacctcg catggcctgc tctgttgccct atagccttac 420
 taagaatccg aaactctccc caaaaagcag gacttagtcc atacaaaatg ctgtatggac 480
 ggcccttccct aaccaatgaa cttgggcttg accgagagac agccaactta gttgcagaca 540
 tcatctcctt agccaaatat caacagggttc ttaaaacatt acagggagcc tgtccccaag 600
 aagagggaaa ggaactattc caccctggtg acatgggtatt agtcaagtcc cttccctcta 660
 attccccatc cctagatata tcctgggaag gaaactaccc agccatttta tetaccctaa 720
 cggcagttaa agtggctgga gcggagtctt ggatacatca cactcaagtc aaaccctgga 780
 tactgccaaa ggaactcaaa aatccatgag acaatgctag ctatttctgt gaacctctag 840
 aggatctgag cctgctcttc aaatgacaac cagggggaaa gtaactaaaa tcgtaaatcc 900
 cctggccctc ccttatcata ttttctctt tactgttctc ttaccocctt tcaactctac 960
 tgcaccccggt ccatgccact gcaccccggt catgccccgt ccatgccagt agctccctt 1020
 agcaagagtt tctatggaga atgcagcgtc ccggaaatat tgatgcccc tttgtatagga 1080
 gtttatctaa gggaaccccc accttactg cccacaccca tatgccccac aactgctata 1140
 actctgccac tctttgcatg catgcaaata ctcattattg gacaggaaaa acgattaatc 1200
 ccagtgtgct tggaggactt ggaggactca cttactcat accagtatgt ctgatggggg 1260
 tggagttcaa gatcaggcaa cagaaaaaca cataaaggaa gtaatctccc aactgacctg 1320
 ggtacatagc acccctggcc cctacaaagg actagatctc tcaaaactac atgaaacct 1380
 ccat acccat actggcctgg taagcctatt taataccacc ctgactgggc tccatgaggt 1440
 ctcgcccaa aaccctacta actgttggt gtgcctcccc ctgcacttta ggccatacat 1500
 ttcaatccct atacctgaac aatggaacaa cttcagcaca gaaataaaca ccaacttctgt 1560
 ttttagtagt cctctttcca atctggaat aacccatacc tcaaacctca cctgtgtaaa 1620
 atttagaggt actatagaca cagccaactc ccaatgcac aggtgggtaa ctcctcccac 1680
 acgaatagtc tgccctacct caggaatatt tttgtctgt ggtacctcag cctatcattg 1740
 tttgaatggc tcttcagaat ctgtgtgctt cctctcattc ttagtgccc ctatgcccac 1800
 ctacactgaa caagatttat acaatcatgt catacctaag ccccgcaaca aaagagtacc 1860
 cattcttctt tttgttattg gagcaggagt gctaggcgga gtagctactg gcattggcgg 1920
 tatcacacc tctactcagt tctactacaa actgtctcaa gaactaaatg gtgacatgga 1980
 atgggtcgct gataccctgg tcaccttgca agatcaactt aactccctag cagcagtagt 2040
 ccttcaaaat cgaagagctt tagacttgct aaccgggaa agcgggggaa cctttttatt 2100
 ttttagaggaa aaatgctgtt gttatgttaa tcaatccgga atcatcacg agaaagttaa 2160
 agaaattcaa ggtcgaatat aacgtagagc aaaggagctg caaaacactg gaccctgggg 2220
 cctcctcagc caatggatgc cctggattct ccccttctta ggacctctag cagctataat 2280
 attgttactc ctctttggac cctgtatctt taacctcctt gtttaagttt tcttttccac 2340
 aatcgaagca gtaaaactac aaatcggtct tcaaatggag cccagatgc agtccatgag 2400
 taaaatctac cacggacccc tggaccggcc tgctagccca tgctctgatg ttaatgacat 2460
 caaaggcacc cctcccgagg aaatctcaac tgcacaacct ctactacgcc ccaattcagc 2520
 aggaagcagt tagagtgggt gttggccaac ctccccacaa gcagttgggt tttcctgttg 2580
 agagggggga ctgagagaca ggaataacta gatttcttag accaactaag aatccctaag 2640
 agtagctggg aaggtgaccg cttccacctt taaccacgg gcttgcaact tagctcacgc 2700
 ccaaccaatc agatactaaa gagagctcac taaaatgcta attaggcaaa aacaggagat 2760
 aaagaaatag ccaatcatct gttg 2784

<210> 5
 <211> 1799
 <212> ADN
 <213> Homo sapiens

<400> 5
 gggattctta gtcggcctag gaaatccagc taatcctgtc tctcagtcce cccactcaac 60
 agggaaaccc aagtgtctgtt ggggaggttg gctgacgacc agtctaactg cttcctgcgc 120
 aattggggca tagtaggggt tgtgcagttg agatttcctc gggaggggtg cgttcgatac 180
 cattacaatt ggagcatggg ctagtagggc ggtccagggg tccacggtag atccttagtca 240
 tggacttcat ctgggggtcc atttgaagaa cgatttgtag ctttacaact ttgattctgg 300
 aagagacaaa ctttaacaagg aggttaaaga tacaggggtc aaagaggagt atcaatatca 360
 gagctgctag agatcctaag aaggggagaa tccagggcat ccattggctg agggagcccc 420
 aggggtctgt gtttttgaag ctctctgtgt ctacgttgta ttcaatctcg aatttcttca 480
 actttctctg tgacaattca ggattgatta acataataac aacattcttc cgctaaaata 540
 acataataac aacattcttc ccctaaaaat aaacagcttc cccctcttct agagggttagc 600
 aagtctaaag ctcttcaatt ttgaaggact actgatgcta ggaagttaag ttgatcttgc 660
 aaggtgacca gggagtcggc aacccattcc atgtcaccat tgagttcttg agatagtttg 720
 tagtagaact gagtagaggt tgtggtaccg ccaatgccag aacctagtcc acctagcact 780
 cctgtctccg taacaaaagg aagaatgagt actcttttgt tgtggggctt aggtacaaca 840
 taattgtata aatcttggtc agtgtaaatg gtcattgggg cactaagaat gagaggaagc 900
 acatagattc tgaagagcca ttcaaacaac gataggctaa ggtaccacag acaaaaaata 960
 ttcttgaggg taggcagact attcgtgtgg gaggagttac ccacctgatg cattgggagt 1020
 tgggtgtgtc tacagtattg ctaaatttta cacaggtgag gtttgaggta tgggttattt 1080
 ccagattgga aacaagaggt cctactaaaa cggaagtggg gtttatttct gtgctgtagt 1140
 tgttccattg ttcaggtaca gggattgaaa tgcattggcc gaaatacagg gggaggcaca 1200
 accaacagtt agtaggggtt tggaccgaga cctcatggag cccagtggag gtggtattaa 1260
 ataggcttac caggcaagta tgggtatgga gggtttcatg tagtttttaag agatctagtc 1320
 ctttgtaggg gctaggggtg ctatgtaccc gggtcagttg ggaggttact tcccttacat 1380
 gtttttctct tgccctgatc tgaactccac cccctcaga cataccagta tgggtgaagt 1440
 aagtccgaca gacagtggct ccaagtcttc caggacaact aggattaatc attttccctg 1500
 tccaataatg agtatttgca tgcattgcaa gattggcaga gttatagcag ttgtggggca 1560
 tatgggtgtg ggcagtgaag gtggagtttc ctttaggtaa actcctattt gatggggcat 1620
 caatatttct ggggaagccgc attcttcata gaaactcttg gtaaggggag ctgctgggtg 1680
 tacagcagca tggagggggg gcagtgagag tgaagggggg taagagaaca gtaagagaa 1740
 aaatatgata agggaggggc atggggattt acgatttttag ttactttcct cacggttgt 1799

<210> 6
 <211> 1489
 <212> ADN
 <213> Homo sapiens

<400> 6
 tgggtgcttg cccgggcaact ctcagtcctg ctgctggatc atctgggttag tggcttctga 60
 ctcaaggagc ctacgtcccc tggggcagtg ggccttacag tgattccctt gacacgaggt 120
 gcatggacga gggggcggtt tatttctatt ttggacaatct tttttaaagt gtccttgtag 180
 accgcaactg aagcaaacc ctttaggcatt ttgatttgcc tagcttttcc cttttccagt 240
 gcctccaaag tccgcttgcc tgagggccat gactaaagcg gtggcctttt ttttatccca 300
 tttgtcccat tctgcctgct catcctgatc tctattataa aaaactgagg ttgccaagtt 360
 caatagggtt tctaagtttt gttccgggac taaggcagac ttttgaagtt ttttccaat 420
 gtctgtagct gactgagtg taaacttata ctttaagatt agttggcctt cagttagtc 480
 agttgacaga gagaggtatg cttcctcaat gcctccgta gtcactccag aaaggcggta 540
 ggattttctt cctttccctg tgttatagtg gacatcattg aataactcac aggttcttt 600
 ctagttttcc ttagtccttc tagcacgcaa gtttagcaat gtctgcggca ccaatctcca 660
 tgttctgatt ctgtgtccca gtgagggctt acactgggaa ctgcctgctg gcctgtgggg 720
 aatcgttctc tttcctctgt tgcgcacctc tcattgacct gactgagata ccagagatcg 780
 ccaactctc aggtctgcagt tacggcgaca cttctgtcat ttgggggttag tgtctgattt 840

agcagtaaca	ttatatctct	ccatatcaga	tcaaaggatt	gtcctaaacc	ttgtaaaaca	900
tcaatatagc	cattagggtt	atctgagaat	ttacctaggt	ctattttaat	ttaaagtctg	960
ggagagaaaa	aggcacatgc	actctggctg	ggccgaattc	tcttccctcc	actgctgtctg	1020
agagagaaaa	aggtacgtgc	actctggctg	ggccgaattc	tcctcccacc	gcttggagggtg	1080
ggcataatcg	gggaatattg	gcattctttg	gttagttgtt	tacccctttg	tctatctctt	1140
tttggaccgt	ttgggttgaa	gggggggtcct	tattatttgg	ggaaggagtc	tgggggatgc	1200
tggggtaggg	aggtagactc	tgagggcttc	ctgtagggca	taaatcacac	tttttacata	1260
attgcgagtt	gtctcttaat	gaaaagaaaag	tttgtacgta	tgacacttca	caccattttgc	1320
cttcttttct	acaaaagagg	tctagctgta	agatgggtgt	ataatttatg	cttccctcag	1380
gatgccaggt	ttctccccct	taaagagtat	atcgttgcca	ggcggtagctg	cagaagaata	1440
tgtctttttt	ttcttagcat	ctgagagtc	aattgggtccc	aattctcca		1489

<210> 7

<211> 1216

<212> ADN

<213> Homo sapiens

<400> 7

taaagataca	gggattgaaa	tgtatggcct	gaagtgcagg	gtcatatagg	tgtgggtggt	60
gaaaatgggg	tttcttttag	aaaaactcct	atacgatggg	tcatcaatat	ttccagggaag	120
ccgcattctc	catagaagct	cttggtaatg	ggagctactg	gtagtacagt	ggcatggagg	180
gggtgcagtg	agagtgaaaag	agggtaaaaag	aacagtaaaag	agaaaaatat	gataaggagg	240
gggttcagtg	agagtgaaaag	gggttaagag	aacagtaaaag	aaaaaaatat	gacaaggagg	300
gccatgagga	tctacgattc	tagttacttt	cctcacgggt	gtcgcttgaa	gagcagggtgc	360
agatcctcta	gaggttcaca	ggaatagcta	gcgtttgtctc	ctggattttc	gggttccttt	420
ggcagtatac	agagtttgac	tcgagtgtga	tgtattcaag	actccactcc	agccacttta	480
accgcagttg	gggtagataa	aatgactggg	taggggtcctt	cccaggatgt	atctaaggat	540
ggggacttag	aaggaaaggga	cttgactaat	accatgtcac	cagggtgcaa	taattacttt	600
ccctcttctc	gggaacaggt	tccctgtaat	gttttaagaa	cttgttgata	tttggccaag	660
gaggtgatgt	ctgcaactaa	gctggccatc	tctcggtcaa	gcacaagggtc	cttgggttagg	720
aaggggccatc	catacagcat	tttgtatggg	ctaagtcctg	ctttttgggg	agagtttttg	780
attcttagta	aggctgtagg	caacagagca	ggccatgcaa	ggtgggtttc	ttgggttagc	840
ttttttaaat	gtcgtttgag	tgtttcattc	attttcttga	cttttcttga	ggattgtggc	900
ctccacgcgc	agtgtaaagt	atattgtatg	cctaattgcct	gggatactcc	ctgggttact	960
gtagccttga	aaacggggcc	attgtcactc	tgtaagcctc	ggggaagtcc	gaatctggga	1020
attatttcat	gaattagtgc	ctttattaca	tcttggtcct	tttctgtcct	acaaaggaag	1080
gcctctgccc	aaccagtga	aatatctacc	cagactagta	gatactgaaa	tccctgagat	1140
ttgggcatgt	gggtaaaatc	tagttgccag	tcttctcctg	agtaatggcc	tgttctttgt	1200
tctcctgaag	gagctt					1216

<210> 8

<211> 976

<212> ADN

<213> Homo sapiens

<400> 8

agtataatg	gaatacttga	aagtaatccc	ctcactccag	gaactagtgc	tgagctggcc	60
aaactaatag	ccctcactcg	ggcactagaa	ttaggagaag	agaaaagggt	aaatatatat	120
acagactata	agtatgctta	cctagtcctt	catgcccatg	cagcaatatg	gagagaaaagg	180
gaattcctaa	cttccaaagg	aacacctatc	aaacatcagg	aagccattag	gatattatta	240
ttggtggtac	agaaaacctaa	agaggtggca	gtcctacact	gctgggttca	tcagaaaaaa	300
aaggaaagg	aaatagaagg	gaactaccaa	gcagatattg	aagccaaaag	agccgcaagg	360
caggaccctc	cattagaaat	gcttatagaa	ggacccttag	tgtggggtaa	ccccctccag	420
gaaagcaatc	cccagtactc	agcaggagaa	ataaaatgga	gaacctcacg	aggacatact	480
ttcctcccc	caggatggct	agccaccaaa	gaaggaaaaa	tgcttttgcc	tgagctaac	540
caatggaaat	tacttaaaac	ccttcaccaa	acctttcact	taggattgat	agcaccatc	600

9

agatggccaa	attattattt	actggatcag	gccttttcaa	aactatcaag	caggtagtca	660
gggcctgtaa	agtgtgccaa	agaaataatc	tcctgcactg	caagccatac	atttcaatcc	720
ctgtatcttt	aacctccttg	ttaagtttgt	ctcttccaga	atcaaagctg	taaaactaca	780
aatggttctt	caaatggagt	ctcagatgca	gtccatgact	aagatatacc	gcagccccct	840
ggagggggcc	tgctagccca	tgctccaatg	ttaatgacat	cgaaggcacc	cctcccgggg	900
aaatctcaac	tgcaacaacc	ctactatgtc	ccaattcagc	aggaagcagt	taaagcggtc	960
atcggccaac	ctcccc					976

<210> 9
 <211> 942
 <212> ADN
 <213> Homo sapiens

<400> 9						
agaggagaac	agcagcataa	gcggctggca	gaggtaggga	aagaccagca	agaagaaaag	60
agagaaaag	aaagagaaaag	tcagagaaaag	agacagagag	aggaagagac	aaagagacag	120
aaagtcaaaag	aggtagtagt	cagaaacaga	gacaaaaaaa	aggagtccga	aagagggaca	180
gacacagaaa	gtcaaaaaaa	aagttaagaa	gaaaggaaaa	gacaaagaag	aagtcgaaga	240
ggagaaaag	agagatagaa	gtagtaaaga	aaaaaacagc	atatccccat	ccttttaaagc	300
cagggtaaat	ttctatctac	ccagccaagg	catattctac	ttatgtggat	cttcaaccca	360
tatctgcctc	tcagacagtt	tgcaagaaat	aatgaaatct	atccttactt	tacaatccca	420
aatagactct	ttggcagcag	tgactctcca	aaactgcaga	ggcctagacc	tcctcactgc	480
tgaaaaagga	ggacactaca	ccttcttagg	ggaagaatgt	tgtttttaca	ctaaccagtc	540
gggtagtagt	tgagatgctg	cccggagttt	acaggaaaaag	gcttctgaaa	tcagacaacg	600
cctttcaaat	tcttatacca	acttctggag	ttaggcaaca	tggttctccc	cctttctagg	660
tcctgtggca	gccatcttgc	tgttactcgc	ccttggggccc	tgtattttta	accttcttgt	720
caaatttggt	tcctctagaa	tcgaggccat	caagctacag	atggtcttac	aaatggaacc	780
ccaaaagagt	tcaactaaca	acttctacog	aggacccctg	gatcaaccca	ctggcacttc	840
ccctggccta	gagagttccc	ctctgaagga	caccgcaact	ggagggccct	tctttgcccc	900
atccagcagg	agtagctaga	gtggtcatcg	gccaaattgc	ca		942

<210> 10
 <211> 1375
 <212> ADN
 <213> Homo sapiens

<400> 10						
ccccaatatt	ctctttctga	tggggaaaaa	tggccacctg	agggaagcac	aaattacaat	60
actatcctgc	agcttgatct	tttctgtaag	agggaaggca	aatggagtga	aataccttat	120
gtccaagctt	tcttttcatt	gaggggagaat	acacaactat	gcaaagcttg	caattttacat	180
cccacaggag	gacccctcag	cttaccoccca	tatcctagcc	tccttatagc	ttcccttcct	240
attgatgata	ctcctcctct	aatctccccct	gcccagaagg	aaataagcaa	agaaatctcc	300
aaaggtccac	aaaaaccccc	gggctatcgg	ttatgtcccc	ttcaagctgt	agggggagggg	360
gaatttggcc	caacccgggt	gcatgtcccc	ttctccctct	ctgattttaa	gcagatcagg	420
cagacctggg	gaagttttca	gatgatcctg	ataggtacat	agatgtccta	cagggtctag	480
ggcaaaccct	tgacctcact	tgagagagacg	tcagtctact	gttagatcaa	accctggcct	540
ttaatgaaaa	gaatgcggct	ttagctgcag	cctgagagtt	tgagataacc	tggtatccta	600
gtcaagtaaa	tgaaagaatg	acagccgaag	aaagggacaa	cttccctact	ggtcagcaag	660
ccatccccag	tatggatccc	cactgggact	ttgactcaga	tcattggggac	tgagatcgta	720
aacatctggt	gatctgtggt	ctggaaggac	taaggagaat	tgggaaaaag	cccatgaatt	780
attcaatgat	atccaccata	accaggggaa	aggaagaaaa	tccttctgcc	ttcctcgagc	840
ggctacaaga	ggccttaaga	aaatatactc	ccctgtcacc	cgaatcactc	gagggtcaat	900
tgattctaaa	agataagttt	attaccctaat	cagccacaga	tatcaggaga	aagctccaaa	960
agcaagccct	gagccctgaa	caaaatctag	agacattatt	aaacctggca	accttgggtgt	1020
tctataatag	ggaccaagag	gaacaggccc	aaaaggaaaa	gcgagatcag	agaaaggccg	1080
cagccttagt	catggccctc	agacaaacaa	accttgggtg	ttcagagag	tcagaaaaatg	1140

```

gagcaggcca atcacctggt acggccttgtt atcagtgcgg ttacttagga cactttaaaa 1200
aagattgtcc aataagaaac aagctgcccc ctcatccgtg tccactatgc cgaggcaatc 1260
actggaaggt gcactgcccc agaggatgaa ggttccctgg gttagaagcc cccaaccaga 1320
tgatccaaca acaggactga gggtgccccg ggcaagcacc agctcatgtc atcac 1375

```

```

<210> 11
<211> 944
<212> ADN
<213> Homo sapiens

```

```

<400> 11
acctaggagg aactgtcttc aggacaggac tatagatgct tcctcccagg cgattaaggg 60
aaaaagacac aatgggtatt cagtaagtga taaggaaact cttgtagaag cagagttagg 120
aaaattgcct aataattggt ctgctcaaat gtgcgagctg ttgcaactca gccaaacctt 180
aaaagtatta cagaatcagg aagaagccat ctataccaat tctaagttaa tatggactga 240
acgagaactt attaatagca aagaataatt gaaatcccaa acttacaagg ttttcaacaa 300
aagcacagtt tgctaaaagt taactgtgta acatgtatta tcctactacc acaaactctc 360
aaatgatttc tcagacagtt tgcaagaaac aatgaaacct atccttactc tacaatcccc 420
aatagactct ttggcagcag tgactctcca aaaccaccaa ggccatagacc tcctcactgc 480
tgagaaagga ggactctgca ccttcttagg ggaagattgt tgtttttaca ctaaccagtc 540
agggatagtg tgagatgcca ccagcggttt acaggaaaag gcttctgaaa tcagacacaa 600
tgcttttcaa accttatagc aacctctgga gttcggcgac tggctttttc cctttctagg 660
tcctgtgaca gccatcttgc tattactcgc ctccggggccc tgtattttta acctcctcgt 720
caaatttggt tcctctagga tcgaggccat caagctacag atggtcttac aaatggaacc 780
ccaaatgagc tcgactaaca acttctactg aggacccctg gaccgacca ctggcccttt 840
aactggctta aagagttttc ctctggagga cactacaact gcaggggccc ttctttgccc 900
catccacagg aagttagcta gagcagtcac cacccaattc ccaa 944

```

```

<210> 12
<211> 963
<212> ADN
<213> Homo sapiens

```

```

<400> 12
tacaggaacc ccataatacg tccttggaac attctattca gctccaactg ctaggagtgg 60
cccatttgtc ctgaaccctc aaatcatggg aatgagaaat gaatttagac tgaccacagc 120
ccttatgagt ttccagctac aggggtgtat agaaccctga taaggagttt tctttgtgtg 180
tggaagatcc ttctatatth gcctcccac caactggaca ggaacttgta ctttagccta 240
catagtacct cctgtgactt atccttttca gaagaggcag tagctgtgcc cattcatgct 300
aagcttcagc cgagagcaat ctactactt cctctattgg ctggtttagg atttactacc 360
acctaggaag tggactcaca gcctagatga aatctctctc caacttactc aaatccagga 420
ccaaatagac tcattagcag ctgtggttct ccgaaccagt gagcactaga tctccaatct 480
cctcactgcc gaaaggggag gaacatgcct ttttctgaac aaggaatgtt gtttttatgt 540
caataaatca ggcatagtga gagatggaat taaatgactt caggatagag cttagcagact 600
acatgggtggg acaaccgaaa ctacctcagg gttctcacag cctgttctcc actggcttct 660
tccattttta ggtcccttcc ttatgattat tctaggagta acctttggcc catgtctttt 720
cagttccttc atcctttcgt ttcttctgta atagaatcaa tgaaactaga aatgttactg 780
cagatggaac ctcatatgac ttcaaccagc acctattatc aaggacccct aaaccagcct 840
gccggcccat acccgagcgt tgacacccaa accacctctc acgaggaaac cttagctaca 900
gaaccccttc tatgccccta ttcagcagga agcaattaga gtggtcatcc tcccacaccc 960
caa 963

```

<210> 13
 <211> 1362
 <212> ADN
 <213> Homo sapiens

<400> 13
 ccacaatatc ctcttccagg aggagaacga tggccacctg aggggaagtat acactataat 60
 accatcctgc aactagatct gttttgtaaa caagaaggca agtggattta ggtaccatat 120
 gttcagacct ttttctcatt aagggatgat aaccacacgat tgtgtaagac atgtaacctg 180
 caccacacag ggagtcctca aattctaccc ccataccacag tcctccccac ggctcctcct 240
 actaatgcca aaccctctct ggcttctaca gcccaaaagg gaacaaataa aagagccttc 300
 agagagccaa gagacccccc tggcccctgg ctatgtcctc ttcaggctgt aggaggggaa 360
 tttggcccaa cccgagtaca tgttcccttt tctctctctg atctaaagca aattaaggca 420
 gacttggtatg aaagtcttca gatgacccca atagatacgt agatggcctg ctgggtctgg 480
 gacaatcttt tgacctttcc tggagagaga tcatgttatt gcttgatcag acctaacctc 540
 taatgagaag aatgctgctt taacaggagc ccgagagttt ggggatacct ggtacctcag 600
 ttaagtaagt gatagaatga catcagaaga gagcagtttc ctactggcca gcaagcagtc 660
 cccagtatgg atccccactg ggacctgac tcggatcatg gggactggag tcacaaacat 720
 ttactgacct gtatcctaga agggtaagg agaactagga aaaagcccat gaactattca 780
 atgatgtcta ctataaccca agggaaggaa gaaaacctta ttgccttctc caaaaggctg 840
 agggaggctt tgagaaaata tactcccctg tcaccagatt ccctcgaagg ccagttaatt 900
 ttaaaggaca aatttattac tcagtcagct gcagacatta ggaaaaagct ccaaaagtta 960
 gccttgggcc gagcaaaatt tggaggcatc attaaacctg gcaacctcag tgttctatca 1020
 tagggacca gaggaacagg ccgaaaagga aaagcaggat aagagaaagg ctgcagattt 1080
 agtcatgccc tcagacaaac cttggcggtt caaagaggag aaaaaatgga gcaggccaat 1140
 caccagcag ggcttattat cagtgcagtt tacaaggaca ctttaaaca gattgtccaa 1200
 agagaaataa gccgccctct caccatgtc cactatgcc aagtgatcac tggaaaggcac 1260
 actgtcccag aggacaaagg ttctctgggc cagaagtccc caaccagatg atccagcaac 1320
 aggatggagg gtgcccgggg caagcaccag ctctgtttgt ca 1362

<210> 14
 <211> 945
 <212> ADN
 <213> Homo sapiens

<400> 14
 ttgcagatca atctcagact gctgtgctag caatgagtga ggcttcgtgg gcatgggacc 60
 ctctgagcca ggcatgggat ataatgtcct tgtgtgccat ttgctaagac tgttgggaata 120
 gcacagtatt aggggtgggag tggcccgatt ttccagggtgc tgtctgtcac cgcttccctt 180
 ggctaggaaa gagaattccc tgacccttg ttcttcccag gtaaggcagt gcctcacctt 240
 gcttcagctc acactcaggt gactgcaccc actgtcctgc cccactgtc ggacaagccc 300
 cagttagatg aacctggtac ctgagttgga aatgcagaaa tcacctgtct tctgcgtcac 360
 tcacactggg agctgtagac tggagctgtt cctatttggc catcttgga ccatctccca 420
 aatagactct ttggcagcag tgactctcca aaaccaccaa ggcctagacc tcctcattgc 480
 tgagaaagga ggactctgca cttctttagg ggaggagtgt tgtttttata ctgaccagtc 540
 agggatggta cgagatgcca cccgatgttt acaggaaaag gcttctgaaa tcacacaaca 600
 cttttcaaac tcttatacca acctctggag ttgggcaaca tggcttctcc cttttctcgg 660
 tcccattgca gccatcttgc tattactcgc cttcaggctg tgtattttta acctccttgt 720
 caaatttgtt tcctctagaa ttgaggccgt caagctacag atggtcttac aaatgggacc 780
 ccaaattgagc tcaactaaca acttctgcca aggacccctg gaccaacctg ctggcccttt 840
 cactggcctt aagagttccc ctctggaggg cactacaact gcagggcccc ttctttgccc 900
 ctatccagca ggaagtagct agagcagtc tccaccaatt cccaa 945

12

<210> 15
 <211> 939
 <212> ADN
 <213> Homo sapiens

<400> 15
 agagctacct tggcaagtac tctaggagta tgggaaaatg aaaacaacaa actcacacac 60
 cattttaaca tacacaatca ggtctgcccc cccagcaagg tataattcttt gtatgtggaa 120
 catcgacctata tatctgcctc cccactaact agacagccac ctgaatctta gtctttctaa 180
 gtcccaacag taacattgcc ccaggaaatc agaccatata agtatccctc aaagctcaag 240
 tctgtcagtg cagagccata caactaatac ccctacttat agggtaagga atggctactg 300
 ctacaggaac cagaatagct agtttggtta cttcattatc ctactaccac acactctcaa 360
 atgatttctc agacagtttg caagaaataa cgaatctat ccttactcta caatcccaaa 420
 tagactcctt ggcagcagtg accctccaaa acggctgagg cctagacctc ctactgcc 480
 agaaaggagg actctgcatt tcttagggg aagagtgttt ttacactaac cagtcaggga 540
 cagtatgaga tgccactcgg agtttacagg aaaaggcttc tgaagtcaga caatgccttt 600
 caaactctat accaaactct ggagttgggc aacatggctt ctcccctttc taggtcccgt 660
 gacagccatc ttgctattat ttgcctttga gccctgtatt tttaatctcc ttttcaaat 720
 tgtttctctt ggatcgaggc catcgagcta cagatggctt tcacaaatgg aacccccaaa 780
 gagctcaact aacaacttct actgaggacc cctggactaa cctgctgacc ctttactctg 840
 cctgaagaat tcccctctgg aggacactac aactgcaggg ctccttcttt gccctatcc 900
 agcaggaagt agctagagct gtcattgcct aattcctaa 939

<210> 16
 <211> 979
 <212> ADN
 <213> Homo sapiens

<400> 16
 agtgataatg gaatacttga aagtaatccc ctactcccc aggaactagt gctcagctgg 60
 cagaactaat agccctcact cgggtactag aatcaggaga aggaanaaagg gtaaatatat 120
 atacagactc taagtgtgct tacctagtcc tccatgcccc tgcagcaata tggagagaaa 180
 gggaattcct aacttccgag ggaacaccta tcaaacatca ggaagccatt aggaattat 240
 tattggctgt acagaaacct aaagagggtg cagttttaca ctgccgggg catcagaaa 300
 gaaaggaaaag ggaaatataa gggagccacc aagttgatat tgaagtcaaa agagccacaa 360
 ggctggaccc tccattagaa atgcttatag gaggaccctc agtatgggg aatcccctcc 420
 gggaagccaa gccccagtag tcagcaggag aatagaata gggaaactca tgaggacata 480
 ctcccctccc ctccagatgg ctagccacca ataaaggaaa aatacttttg cctgcagcta 540
 accaatagaa attacttaaa acccttcac cagccctttt caaaactatc aagcagatag 600
 atgagatggc caaattatta tttactggac caggcctttt caaaactatc aagcagatag 660
 tcagggcctg taaagtctgc caaagaaata atcccctgca ctgcaggcca tacatttcaa 720
 tccctgtatc tttaacctcc ttcttaaaatt tgtctcttcc agaatcaaa ctgtaaaatt 780
 acaaatagtt cttcaaatgg agccacagat gcagtcocat actaagatcc accacagacc 840
 cctggaccag cctgctagcc catgctccaa tgtaaatgac atcgaaggca ccccctctg 900
 aggaaatctc aactgcacaa cccctactac gcccgaattc agcagaaaagc agttagagt 960
 gtcacagcc aacctcccc 979

<210> 17
 <211> 1774
 <212> ADN
 <213> Homo sapiens

<400> 17
 catgctggta aaggaccgct agaatccagc agccaggacc actttctttg tggtaagaa 60
 aggtgggaaa acaggtgcag gactgctaca ctggtaagca taactaatcc gataagcaga 120
 ggtccatggg tggttacgca ccctggaaa gaataagcat taggactata gaggacactc 180

taggactaat	gctcatcgga	aaatgactag	gggtactggc	atccctatgt	tcttttttca	240
gatgggaaat	gttcccccca	aggcagaaat	gcccctaaga	tgtattctgg	agaaatggga	300
ccaatctgac	catcagacac	taagaaagaa	atgacttata	ttcttctgca	gtaccacctg	360
gccacaatat	cttcttcaag	gggcagaaac	ctggcctcct	gaggggaagta	taaattataa	420
caccatctta	cagctagacc	tcttttgtag	aaaagaaggc	aaatggagtg	aagtgccata	480
tgtacaaact	ttcttttcat	taagagataa	ctcccaatta	tgtaaaaagt	gtgatttatg	540
ccctacagga	agccctcaga	gtctacctcc	cgaccccagc	aagaccccaa	ctccttctcc	600
aactaataag	gacccccctt	caacccaaat	ggtccaaaag	gagatagaca	aaggggtaaa	660
caatgaacca	aagagtgcc	atattacacg	attatactcg	ctccaagcag	tgggaggaga	720
at ttggccca	gccagcgtgc	atgtaccttt	ttctctctca	gatttaaagc	aaattaaaat	780
agacctaggt	aaattctcag	ataaccctga	tggctatatt	gatgttttac	aagggttagg	840
acaatccttt	gatctgacat	ggagagatat	aatgttactg	ctaaatcaga	cactaaccctc	900
aaatgaaaaa	agtgtgtgcc	taacagcagc	ctgagagtgt	ggcgaactct	ggtatctcag	960
tcagggtcaat	gataggatga	caacagatga	aagagaatga	ttccccacag	gccagcaggc	1020
agttcccaggt	gtagaccctc	attaggacac	agaatcagaa	cttggagatt	ggtgccacag	1080
acatttgcta	acttgctgtc	tagaaggact	aaggaaaact	aggaagaagc	ccatgaatta	1140
ttcaatgatg	tcccctataa	cacagggaaa	ggaagaaaat	cctactgcc	ttctggagag	1200
actaagggaa	ggattgagga	agcatacctc	cctgtcacct	gactctatta	aaggccaact	1260
aatcttaaa	gataagttaa	tcactcagtc	agctgcagag	attaagaaaa	aacttcaaaa	1320
gtatgcctta	ggcccagagc	aaaacttaga	aaccctactg	aacttggcaa	cctcagtttt	1380
ttataataga	gatcaggaag	agcaggggaa	tgggacaaat	gggataaaaa	aaaaaaaaaa	1440
aggtgactgc	tttagtctgt	gccctcaggc	aaatggactt	tggaggctcc	agaaaaggga	1500
aaagctgagc	aaattgaatg	cctaacaggg	cttgcttcta	gtgtggtcta	caaggacact	1560
ttaaaaaaga	ttgtccaagt	agaaacaagc	tgcccccttg	tccatgcccc	ttatgtcaag	1620
ggaatcactg	gaaggccac	tgcccagga	gatgaaggtc	ctctgagtca	gaagccacta	1680
accagataat	ccagcagcag	gactgaggat	gcccagggca	agcgccagcc	catgccatca	1740
ccctcacaga	gccttgggta	tgcttgacca	ttga			1774

<210> 18
 <211> 938
 <212> ADN
 <213> Homo sapiens

<400> 18						
tgttaggaaga	actcccttca	ggacaggaca	atagatgggt	cctcccaggt	gattaaggaa	60
aaaagacaca	gtattcagta	agtgataagg	aaactcttgt	agaagcagag	ttagaaaaat	120
tgccataata	ttggtctgct	caaattgtgtg	agttgtttgc	actcagccaa	atcttaaaagt	180
acttacagaa	tcaggaagca	gccatctata	ccaattctaa	gttaatatgg	actaaacgag	240
gtttttattag	tagcaaaagaa	aaattaaaaat	cccaaactta	caagggtttc	aactaaagtt	300
tgccaaaagt	taacagtgt	acatgtatta	tcctactatc	acacactctc	aaaggatttc	360
tcagacagtt	tgcaagaaat	aacgtaatct	atccttactc	tacagtccca	aatagactct	420
ttggtagcag	tgactctcca	aaactgccga	ggtctagacc	tcctcaatgc	tgagaaagga	480
gaactctgca	ccttcttagg	ggaagagtgc	tgtttttaca	ctaaccagtc	agggatagta	540
tgagatactg	cctgacgttt	acaggaaaag	gcttctgaaa	tcagacaacg	cctttcaagc	600
tcttatacca	acctctggag	ttgggcaaca	tggcttctcc	ccttgctagg	tctgtggca	660
gccatcttgc	tattacttgc	cttcggggccc	tgtattttta	acctccttgt	caaatttgtt	720
tcctctagga	tcaaggccat	caagctacag	atggtcttac	aaatggaacc	ccaaatgagc	780
tcaactaaca	acttctactg	aggacacctg	gactgaccca	ctggcccttt	cactggccta	840
aagagttccc	ttctggagga	cactacaact	gcaggggccc	gtcttcaccc	ctatccagca	900
ggaagtagct	agatcagtc	ttgcccattt	cccaacag			938

<210> 19
 <211> 1308
 <212> ADN
 <213> Homo sapiens

<400> 19

```

gatgcttgcc ccaggcacc ctagtccctgt tgttggatca tctggtcggg ggcttctggc 60
ccaaagaacc tttgtccctct gaggcagctg accttccagt gattgcctca gcattgtgga 120
catgggcaag ggggcagctt gtttctcact ggacaatctt ttttaagggt tccttccaaa 180
ccacactggg aacaagccct accaggtgat tggcctgctc tatcttctgt cctctctgaa 240
ccaccaagggt ttgtctgtct gaggggtcatg actaaggctg tggcctttct ctgatcttgc 300
ttttcctttt tggcctgttc ctcttgggtac ctattataga aactgaggtg tgccaggttt 360
aacaatggct ccagattttg ttcagggcac aggggtcatt ttggagcttt ctctgatata 420
ctgcagctga ttgggtaata aacttatctt ttaggatcaa ttgactctca agagagttgg 480
gtgacagggg agtatatttc cttgaggcct cccatagccg ctctaggaag gcagaaggat 540
tttcttcctt tccctgagtt ataaaagaca tcattgaaca actcatggac tttttcccaa 600
ttctccgtag tcttctctaga acacaggtca gcagatgttt acgactccag tccccatgat 660
ctgagcttag acaccagtg ggatccatac tggggatggc ctgctgactg gtagggaatt 720
tgtccctttc tttggctgtc attctatcat ttacttgact aagataccaa gtatctccaa 780
attctcaggt tgcagctaaa gctgcattct tttcattaaa ggccagggtt tgatctaata 840
gcatagacatc tctccaagtg aggtcaaagg tttgccctag atccatagga catcagagaa 900
ggagaagggg acatacacct gagttagcca aattcccctc cctctacagc ttgaagggga 960
cataagcaat agcctgggga tttttgtggt cctttggaga tttctttgct tgtttccttc 1020
tggttggggg agattagagg aggttatca gtaataggaa ggggagctat agggaggcta 1080
ggatatgggg gtaagctgag aggtcatctt gtgggatgta aattgcaagc tttgcatagt 1140
tgtggatttt ccttacaatg aaaataaagc ttggacataa ggtatttcac tccatttgcc 1200
ttccctctta cagaaaagggt caagctgcag gatagtactg taattttatac ttccttcagg 1260
tggccatttc tcccatcag agagagaata ctggggctgg gccatagt 1308

```

<210> 20

<211> 711

<212> ADN

<213> Homo sapiens

<400> 20

```

actgagagac aggactagct ggatttccta ggccgactaa gaatccctaa gcctagctgg 60
gaaggtgacc acgtccacct ttaaaccagg ggcttgcaac ttagctcaca cctgaccaat 120
cagagagctc actaaaatgc taattaggca aagacaggag gtaaagaaat agccaatcat 180
ctattgcctg agagcacagc aggagggaca acaatcggga tataaaccca ggcatctgag 240
ctggcaacag cagccccctt ttgggtccct tccctttgta tgggagctgt tttcatgcta 300
tttcaactca ttaaactctg caactgcaact cttctgggtcc atgtttctta cggctcgagc 360
tgagcttttg ctcaccgtcc accactgctg tttgccacca ccgcagacct gccgctgact 420
cccctccctc tggatcctgc aggggtgtccg ctgtgctcct gatccagcga ggcgcccatt 480
gccgtcccca attgggctaa aggccttgcca ttgttccctg acggctaagt gcctgggttt 540
gttctaattg agctgaacac tagtcactgg gttccatggg tctcttctgt gaccacaggc 600
ttctaataga actataacac ttaccacatg gcccaagatt ccattccttg gaatccgtga 660
ggccaagaac tccaggtcag agaatacgag gcttgccacc atcttggaag c 711

```

<210> 21

<211> 711

<212> ADN

<213> Homo sapiens

<400> 21

```

actgagagac aggactagct ggatttccta ggctgactaa gaatccctaa gcctagctgg 60
gaaggtgacc acatccacct ttaaaccagg ggcttgcaac ttagctcaca cctgaccaat 120
cagagagctc actaaaatgc taattaggca aagacaggag gtaaagaaat agccaatcat 180
ctattgcctg agagcacagc aggagggaca atgatcggga tataaaccca agtcttcgag 240
ccggcaacgg caacccccct tgggtccctt ccctttgtat gggagctctg ttttcatgct 300
atttcaactc attaaatctt gcaactgcac tcttctggtc catgtttctt acggcttgag 360
ctgagctttc gctcgccatc caccactgct gtttgccgcc accgcagacc cgccgctgac 420

```

15

```

tcccatccct ctggatcatg cagggtgtcc gctgtgtctc tgatccagcg aggcacccat 480
tgccgctccc aatcgggcta aaggcttgcc attgttcctg catggctaag tgcctgggtt 540
catcctaatt gagctgaaca ctagtactg ggttccatgg ttctcttctg tgacccacag 600
cttctaatag agctataaca ctcaccgcat ggccaaggt tccattcctt gaatccataa 660
ggccaagaac cccagggtcag agaacacgag gcttgccacc atcttgggag c 711

```

<210> 22
 <211> 2055
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(2055)

<400> 22

```

ccc aag aca gcc aac tta gtt gca gac atc acc tcc tta gcc aaa tat 48
Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr
  1          5          10          15

caa caa gtt ctt aaa aca tta caa gga acc tat ccc tga gaa gag gga 96
Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro Glu Glu Gly
      20          25          30

aaa gaa cta ttc cac cct tgt gac atg gta tta gtc aag tcc ctt ccc 144
Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys Ser Leu Pro
      35          40          45

tct aat tcc cca tcc cta gat aca tcc tgg gaa gga ccc tac cca gtc 192
Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro Tyr Pro Val
      50          55          60

att tta tct acc cca act gcg gtt aaa gtg gct gga gtg gag tct tgg 240
Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val Glu Ser Trp
      65          70          75          80

ata cat cac act tga gtc aaa tcc tgg ata ctg cca aag gaa cct gaa 288
Ile His His Thr Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu
      85          90          95

aat cca gga gac aac gct agc tat tcc tgt gaa cct cta gag gat ttg 336
Asn Pro Gly Asp Asn Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu
      100          105          110

cgc ctg ctc ttc aaa caa caa cca gga gga aag taa cta aaa tca taa 384
Arg Leu Leu Phe Lys Gln Gln Pro Gly Gly Lys Leu Lys Ser
      115          120          125

atc ccc atg gcc ctc cct tat cat att ttt ctc ttt act gtt ctt tta 432
Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu
      130          135          140

ccc tct ttc act ctc act gca ccc cct cca tgc cgc tgt atg acc agt 480
Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser
      145          150          155          160

```

16

agc tcc cct tac caa gag ttt cta tgg aga atg cag cgt ccc gga aat	528
Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn	
165 170 175	
att gat gcc cca tcg tat agg agt ctt tct aag gga acc ccc acc ttc	576
Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe	
180 185 190	
act gcc cac acc cat atg ccc cgc aac tgc tat cac tct gcc act ctt	624
Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu	
195 200 205	
tgc atg cat gca aat act cat tat tgg aca gga aaa atg att aat cct	672
Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro	
210 215 220	
agt tgt cct gga gga ctt gga gtc act gtc tgt tgg act tac ttc acc	720
Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr	
225 230 235 240	
caa act ggt atg tct gat ggg ggt gga gtt caa gat cag gca aga gaa	768
Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu	
245 250 255	
aaa cat gta aaa gaa gta atc tcc caa ctc acc cgg gta cat ggc acc	816
Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr	
260 265 270	
tct agc ccc tac aaa gga cta gat ctc tca aaa cta cat gaa acc ctc	864
Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu	
275 280 285	
cgt acc cat act cgc ctg gta agc cta ttt aat acc acc ctc act ggg	912
Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly	
290 295 300	
ctc cat gag gtc tcg gcc caa aac cct act aac tgt tgg ata tgc ctc	960
Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu	
305 310 315 320	
ccc ctg aac ttc agg cca tat gtt tca atc cct gta cct gaa caa tgg	1008
Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp	
325 330 335	
aac aac ttc agc aca gaa ata aac acc act tcc gtt tta gta gga cct	1056
Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro	
340 345 350	
ctt gtt tcc aat ctg gaa ata acc cat acc tca aac ctc acc tgt gta	1104
Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val	
355 360 365	
aaa ttt agc aat act aca tac aca acc aac tcc caa tgc atc agg tgg	1152
Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp	
370 375 380	

gta act cct ccc aca caa ata gtc tgc cta ccc tca gga ata ttt ttt	1200
Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe	
385 390 395 400	
gtc tgt ggt acc tca gcc tat cgt tgt ttg aat ggc tct tca gaa tct	1248
Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser	
405 410 415	
atg tgc ttc ctc tca ttc tta gtg ccc cct atg acc atc tac act gaa	1296
Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu	
420 425 430	
caa gat tta tac agt tat gtc ata tct aag ccc cgc aac aaa aga gta	1344
Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val	
435 440 445	
ccc att ctt cct ttt gtt ata gga gca gga gtg cta ggt gca cta ggt	1392
Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly	
450 455 460	
act ggc att ggc ggt atc aca acc tct act cag ttc tac tac aaa cta	1440
Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu	
465 470 475 480	
tct caa gaa cta aat ggg gac atg gaa cgg gtc gcc gac tcc ctg gtc	1488
Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val	
485 490 495	
acc ttg caa gat caa ctt aac tcc cta gca gca gta gtc ctt caa aat	1536
Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn	
500 505 510	
cga aga gct tta gac ttg cta acc gct gaa aga ggg gga acc tgt tta	1584
Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu	
515 520 525	
ttt tta ggg gaa gaa tgc tgt tat tat gtt aat caa tcc gga atc gtc	1632
Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val	
530 535 540	
act gag aaa gtt aaa gaa att cga gat cga ata caa cgt aga gca gag	1680
Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu	
545 550 555 560	
gag ctt cga aac act gga ccc tgg ggc ctc ctc agc caa tgg atg ccc	1728
Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro	
565 570 575	
tgg att ctc ccc ttc tta gga cct cta gca gct ata ata ttg cta ctc	1776
Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu	
580 585 590	
ctc ttt gga ccc tgt atc ttt aac ctc ctt gtt aac ttt gtc tct tcc	1824
Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser	
595 600 605	

18

aga atc gaa gct gta aaa cta caa atg gag ccc aag atg cag tcc aag 1872
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys
 610 615 620

act aag atc tac cgc aga ccc ctg gac cgg cct gct agc cca cga tct 1920
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser
 625 630 635 640

gat gtt aat gac atc aaa ggc acc cct cct gag gaa atc tca gct gca 1968
 Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala
 645 650 655

caa cct cta cta cgc ccc aat tca gca gga agc agt tag agc ggt cgt 2016
 Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser Ser Gly Arg
 660 665 670

cgg cca acc tcc cca aca gca ctt agg ttt tcc tgt tga 2055
 Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys
 675 680 685

<210> 23
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 23
 Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr
 1 5 10 15
 Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro
 20 25

<210> 24
 <211> 55
 <212> PRT
 <213> Homo sapiens

<400> 24
 Glu Glu Gly Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys
 1 5 10 15
 Ser Leu Pro Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro
 20 25 30
 Tyr Pro Val Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val
 35 40 45
 Glu Ser Trp Ile His His Thr
 50 55

<210> 25
 <211> 38
 <212> PRT
 <213> Homo sapiens

19

<400> 25

Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu Asn Pro Gly Asp Asn
 1 5 10 15

Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu Arg Leu Leu Phe Lys
 20 25 30

Gln Gln Pro Gly Gly Lys
 35

<210> 26

<211> 540

<212> PRT

<213> Homo sapiens

<400> 26

Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu
 1 5 10 15

Pro Ser Phe Thr Leu Thr Ala Pro Pro Cys Arg Cys Met Thr Ser
 20 25 30

Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn
 35 40 45

Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe
 50 55 60

Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu
 65 70 75 80

Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro
 85 90 95

Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr
 100 105 110

Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu
 115 120 125

Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr
 130 135 140

Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu
 145 150 155 160

Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly
 165 170 175

Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu
 180 185 190

Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp
 195 200 205

Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro
 210 215 220

Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val
 225 230 235 240
 Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp
 245 250 255
 Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe
 260 265 270
 Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser
 275 280 285
 Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu
 290 295 300
 Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val
 305 310 315 320
 Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly
 325 330 335
 Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu
 340 345 350
 Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val
 355 360 365
 Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn
 370 375 380
 Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu
 385 390 395 400
 Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val
 405 410 415
 Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu
 420 425 430
 Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro
 435 440 445
 Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu
 450 455 460
 Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser
 465 470 475 480
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys
 485 490 495
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser
 500 505 510
 Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala
 515 520 525

21

Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser
 530 535 540

<210> 27
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 27
 Ser Gly Arg Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys
 1 5 10 15

<210> 28
 <211> 1080
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1080)

<400> 28
 acc tct ttt gta gaa aag gca aat gga gtg aag tgc cat aag tac aaa 48
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
 1 5 10 15

ctt tct ttt cat taa gag aca act cac aat tat gta aaa agt gtg att 96
 Leu Ser Phe His Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile
 20 25 30

tat gcc cta cag gaa gcc ttc aga gtc tac ctc cct atc cca gca tcc 144
 Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser
 35 40 45

ccg act cct tcc cca act aat aag gac ccc cct tca acc caa atg gtc 192
 Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val
 50 55 60

caa aag gag ata gac aaa agg gta aac agt gaa cca aag agt gcc aat 240
 Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn
 65 70 75 80

att ccc caa tta tga ccc ctc caa gca gtg gga gga aga gaa ttc ggc 288
 Ile Pro Gln Leu Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly
 85 90 95

cca gcc aga gtg cat gtg cct ttt tct ctc cca gac tta aag caa ata 336
 Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile
 100 105 110

aaa aca gac tta ggt aaa ttc tca gat aac cct gat ggc tat att gat 384
 Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp
 115 120 125

22

gtt tta caa ggg tta gga caa ttc ttt gat ctg aca tgg aga gat ata	432
Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile	
130 135 140	
atg tca ctg cta aat cag aca cta acc cca aat gag aga agt gcc acc	480
Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr	
145 150 155 160	
ata act gca gcc tga gag ttt ggc gat ctc tgg tat ctc agt cag gtc	528
Ile Thr Ala Ala Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val	
165 170 175	
aat gat agg atg aca aca gag gaa aga gaa tga ttc ccc aca ggc cag	576
Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Phe Pro Thr Gly Gln	
180 185 190	
cag gca gtt ccc agt cta gac cct cat tgg gac aca gaa tca gaa cat	624
Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu His	
195 200 205	
gga gat tgg tgc tgc aga cat ttg cta act tgt gtg cta gaa gga cta	672
Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly Leu	
210 215 220	
agg aaa act agg aag aag tct atg aat tac tca atg atg tcc acc ata	720
Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr Ile	
225 230 235 240	
aca cag gga agg gaa gaa aat cct act gcc ttt ctg gag aga cta agg	768
Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu Arg	
245 250 255	
gag gca ttg agg aag cgt gcc tct ctg tca cct gac tct tct gaa ggc	816
Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu Gly	
260 265 270	
caa cta atc tta aag cgt aag ttt atc act cag tca gct gca gac att	864
Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp Ile	
275 280 285	
aga aaa aaa ctt caa aag tct gcc gta ggc ccg gag caa aac tta gaa	912
Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu Glu	
290 295 300	
acc cta ttg aac ttg gca acc tcg gtt ttt tat aat aga gat cag gag	960
Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln Glu	
305 310 315 320	
gag cag gcg gaa cag gac aaa cgg gat taa aaa aaa ggc cac cgc ttt	1008
Glu Gln Ala Glu Gln Asp Lys Arg Asp Lys Lys Gly His Arg Phe	
325 330 335	
agt cat gac cct cag gca agt gga ctt tgg agg ctc tgg aaa agg gaa	1056
Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg Glu	
340 345 350	

23

aag ctg ggc aaa ttg aat gcc taa
 Lys Leu Gly Lys Leu Asn Ala
 355 360

1080

<210> 29
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 29
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
 1 5 10 15
 Leu Ser Phe His
 20

<210> 30
 <211> 63
 <212> PRT
 <213> Homo sapiens

<400> 30
 Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile Tyr Ala Leu Gln Glu
 1 5 10 15
 Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser Pro Thr Pro Ser Pro
 20 25 30
 Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val Gln Lys Glu Ile Asp
 35 40 45
 Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn Ile Pro Gln Leu
 50 55 60

<210> 31
 <211> 79
 <212> PRT
 <213> Homo sapiens

<400> 31
 Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly Pro Ala Arg Val His
 1 5 10 15
 Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile Lys Thr Asp Leu Gly
 20 25 30
 Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu
 35 40 45
 Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile Met Ser Leu Leu Asn
 50 55 60
 Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr Ile Thr Ala Ala
 65 70 75

24

<210> 32
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 32
 Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val Asn Asp Arg Met Thr
 1 5 10 15
 Thr Glu Glu Arg Glu
 20

<210> 33
 <211> 142
 <212> PRT
 <213> Homo sapiens

<400> 33
 Phe Pro Thr Gly Gln Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp
 1 5 10 15
 Thr Glu Ser Glu His Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys
 20 25 30
 Val Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser
 35 40 45
 Met Met Ser Thr Ile Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe
 50 55 60
 Leu Glu Arg Leu Arg Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro
 65 70 75 80
 Asp Ser Ser Glu Gly Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln
 85 90 95
 Ser Ala Ala Asp Ile Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro
 100 105 110
 Glu Gln Asn Leu Glu Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr
 115 120 125
 Asn Arg Asp Gln Glu Glu Gln Ala Glu Gln Asp Lys Arg Asp
 130 135 140

<210> 34
 <211> 29
 <212> PRT
 <213> Homo sapiens

<400> 34
 Lys Lys Gly His Arg Phe Ser His Asp Pro Gln Ala Ser Gly Leu Trp
 1 5 10 15

25

Arg Leu Trp Lys Arg Glu Lys Leu Gly Lys Leu Asn Ala
 20 25

<210> 35

<211> 685

<212> PRT

<213> Homo sapiens

<400> 35

Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr
 1 5 10 15

Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro Xaa Glu Glu Gly
 20 25 30

Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys Ser Leu Pro
 35 40 45

Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro Tyr Pro Val
 50 55 60

Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val Glu Ser Trp
 65 70 75 80

Ile His His Thr Xaa Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu
 85 90 95

Asn Pro Gly Asp Asn Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu
 100 105 110

Arg Leu Leu Phe Lys Gln Gln Pro Gly Gly Lys Xaa Leu Lys Ser Xaa
 115 120 125

Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu
 130 135 140

Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser
 145 150 155 160

Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn
 165 170 175

Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe
 180 185 190

Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu
 195 200 205

Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro
 210 215 220

Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr
 225 230 235 240

Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu
 245 250 255

Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr
 260 265 270
 Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu
 275 280 285
 Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly
 290 295 300
 Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu
 305 310 315 320
 Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp
 325 330 335
 Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro
 340 345 350
 Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val
 355 360 365
 Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp
 370 375 380
 Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe
 385 390 395 400
 Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser
 405 410 415
 Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu
 420 425 430
 Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val
 435 440 445
 Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly
 450 455 460
 Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu
 465 470 475 480
 Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val
 485 490 495
 Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn
 500 505 510
 Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu
 515 520 525
 Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val
 530 535 540
 Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu
 545 550 555 560

Glu	Leu	Arg	Asn	Thr 565	Gly	Pro	Trp	Gly	Leu 570	Leu	Ser	Gln	Trp	Met 575	Pro
Trp	Ile	Leu	Pro 580	Phe	Leu	Gly	Pro	Leu 585	Ala	Ala	Ile	Ile	Leu 590	Leu	Leu
Leu	Phe	Gly 595	Pro	Cys	Ile	Phe	Asn 600	Leu	Leu	Val	Asn	Phe 605	Val	Ser	Ser
Arg	Ile 610	Glu	Ala	Val	Lys	Leu 615	Gln	Met	Glu	Pro	Lys 620	Met	Gln	Ser	Lys
Thr 625	Lys	Ile	Tyr	Arg	Arg 630	Pro	Leu	Asp	Arg	Pro 635	Ala	Ser	Pro	Arg	Ser 640
Asp	Val	Asn	Asp	Ile 645	Lys	Gly	Thr	Pro	Pro 650	Glu	Glu	Ile	Ser	Ala 655	Ala
Gln	Pro	Leu	Leu 660	Arg	Pro	Asn	Ser	Ala 665	Gly	Ser	Ser	Xaa	Ser 670	Gly	Arg
Arg	Pro	Thr 675	Ser	Pro	Thr	Ala	Leu 680	Arg	Phe	Ser	Cys	Xaa 685			

```
<210> 36
<211> 360
<212> PRT
<213> Homo sapiens
```

<400> 36															
Thr	Ser	Phe	Val	Glu	Lys	Ala	Asn	Gly	Val	Lys	Cys	His	Lys	Tyr	Lys
1				5					10					15	
Leu	Ser	Phe	His	Xaa	Glu	Thr	Thr	His	Asn	Tyr	Val	Lys	Ser	Val	Ile
			20					25					30		
Tyr	Ala	Leu	Gln	Glu	Ala	Phe	Arg	Val	Tyr	Leu	Pro	Ile	Pro	Ala	Ser
		35					40					45			
Pro	Thr	Pro	Ser	Pro	Thr	Asn	Lys	Asp	Pro	Pro	Ser	Thr	Gln	Met	Val
	50					55					60				
Gln	Lys	Glu	Ile	Asp	Lys	Arg	Val	Asn	Ser	Glu	Pro	Lys	Ser	Ala	Asn
65					70					75					80
Ile	Pro	Gln	Leu	Xaa	Pro	Leu	Gln	Ala	Val	Gly	Gly	Arg	Glu	Phe	Gly
				85					90					95	
Pro	Ala	Arg	Val	His	Val	Pro	Phe	Ser	Leu	Pro	Asp	Leu	Lys	Gln	Ile
			100					105					110		
Lys	Thr	Asp	Leu	Gly	Lys	Phe	Ser	Asp	Asn	Pro	Asp	Gly	Tyr	Ile	Asp
		115					120					125			
Val	Leu	Gln	Gly	Leu	Gly	Gln	Phe	Phe	Asp	Leu	Thr	Trp	Arg	Asp	Ile
	130					135					140				

28

Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr
 145 150 155 160

Ile Thr Ala Ala Xaa Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val
 165 170 175

Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Xaa Phe Pro Thr Gly Gln
 180 185 190

Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu His
 195 200 205

Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly Leu
 210 215 220

Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr Ile
 225 230 235 240

Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu Arg
 245 250 255

Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu Gly
 260 265 270

Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp Ile
 275 280 285

Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu Glu
 290 295 300

Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln Glu
 305 310 315 320

Glu Gln Ala Glu Gln Asp Lys Arg Asp Xaa Lys Lys Gly His Arg Phe
 325 330 335

Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg Glu
 340 345 350

Lys Leu Gly Lys Leu Asn Ala Xaa
 355 360

<210> 37
 <211> 26
 <212> ADN
 <213> Homo sapiens

<400> 37
 ggacataga ggacactcca ggacta

26

<210> 38
 <211> 25
 <212> ADN
 <213> Homo sapiens

<400> 38
cctcagtcct gctgctggat catct 25

<210> 39
<211> 27
<212> ADN
<213> Homo sapiens

<400> 39
cctccaagca gtgggaggaa gagaatt 27

<210> 40
<211> 28
<212> ADN
<213> Homo sapiens

<400> 40
ccttcctgt gttattgtg acatcatt 28

<210> 41
<211> 30
<212> ADN
<213> Homo sapiens

<400> 41
ggaagaagtc tatgaattat tcaatgatgt 30

<210> 42
<211> 27
<212> ADN
<213> Homo sapiens

<400> 42
gggacacaga atcagaacat ggagatt 27

<210> 43
<211> 27
<212> ADN
<213> Homo sapiens

<400> 43
gccttcagaa gagtcaggtg acagaga 27

<210> 44
<211> 25
<212> ADN
<213> Homo sapiens

<400> 44
gagcctccaa agtccacttg cctga 25

<210> 45
<211> 29
<212> ADN
<213> Homo sapiens

<400> 45
gatttcagta tctactagtc tgggtagat 29

<210> 46
<211> 27
<212> ADN
<213> Homo sapiens

<400> 46
ctaggaaatc cagctagtcc tgtctca 27

<210> 47
<211> 28
<212> ADN
<213> Homo sapiens

<400> 47
ccaagacagc caacttagtt gcagacat 28

<210> 48
<211> 28
<212> ADN
<213> Homo sapiens

<400> 48
ggacgctgca ttctccatag aaactctt 28

<210> 49
<211> 29
<212> ADN
<213> Homo sapiens

<400> 49
gcaatactac atacacaacc aactcccaa 29

<210> 50
<211> 26
<212> ADN
<213> Homo sapiens

<400> 50
gggggaggca tatccaacag ttagta 26

31

<210> 51
<211> 30
<212> ADN
<213> Homo sapiens

<400> 51
ccatctacac tgaacaagat ttatacactt

30

<210> 52
<211> 28
<212> ADN
<213> Homo sapiens

<400> 52
aatgccagta cctagtgcac ctagcact

28

<210> 53
<211> 31
<212> ADN
<213> Homo sapiens

<400> 53
cgaataacaac gtagagcaga ggagcttcga a

31

<210> 54
<211> 28
<212> ADN
<213> Homo sapiens

<400> 54
agcccaagat gcagtccaag actaagat

28

<210> 55
<211> 27
<212> ADN
<213> Homo sapiens

<400> 55
gcgtagtaga gggtgtgcag ctgagat

27

<210> 56
<211> 27
<212> ADN
<213> Homo sapiens

<400> 56
cccttaccaa gagtttctat ggagaat

27

<210> 57
<211> 27
<212> ADN
<213> Homo sapiens

32

<400> 57
accgctctaa ctgcttcctg ctgaatt

27

<210> 58
<211> 420
<212> PRT
<213> Homo sapiens

<400> 58
Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
1 5 10 15
Leu Ser Phe His Xaa Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile
20 25 30
Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Leu Pro Ala
35 40 45
Ser Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met
50 55 60
Val Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala
65 70 75 80
Asn Ile Pro Gln Leu Xaa Pro Leu Gln Ala Val Gly Gly Arg Glu Phe
85 90 95
Gly Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln
100 105 110
Ile Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile
115 120 125
Asp Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp
130 135 140
Ile Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala
145 150 155 160
Thr Ile Thr Ala Ala Xaa Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln
165 170 175
Val Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Xaa Phe Pro Thr Gly
180 185 190
Gln Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu
195 200 205
His Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly
210 215 220
Leu Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr
225 230 235 240
Ile Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu
245 250 255

33

Arg Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu
 260 265 270
 Gly Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp
 275 280 285
 Ile Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu
 290 295 300
 Glu Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln
 305 310 315 320
 Glu Glu Gln Ala Glu Gln Asp Lys Arg Asp Xaa Lys Lys Gly His Arg
 325 330 335
 Phe Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg
 340 345 350
 Glu Lys Leu Gly Lys Leu Asn Ala Xaa Xaa Gly Leu Leu Pro Val Arg
 355 360 365
 Ser Thr Arg Thr Leu Xaa Lys Arg Leu Ser Lys Xaa Lys Xaa Ala Ala
 370 375 380
 Pro Ser Ser Met Pro Leu Ile Ser Arg Glu Ser Leu Glu Gly Pro Leu
 385 390 395 400
 Pro Gln Gly Thr Lys Val Leu Xaa Val Arg Ser His Xaa Pro Asp Ser
 405 410 415
 Ser Ser Arg Thr
 420

<210> 59
 <211> 32
 <212> ADN
 <213> Homo sapiens

<400> 59
 taaactacaa atggttcttc aaatggagcc ca

32

<210> 60
 <211> 32
 <212> ADN
 <213> Homo sapiens

<400> 60
 gatgcagtc aagatgcagt ccatgactaa ga

32

<210> 61
 <211> 1740
 <212> ADN
 <213> Homo sapiens

<400> 61

```

aggttggctg acaaccgctc ttaactgctt catgctgaat tggggcatag taggggtcgt 60
gcagttgaga tttccttggg aggggtgcct tcaatgtcat caacattgga gcatgggcta 120
gcaggccagt ccaggggtcc gcggtagatc ttagtcatgg actgcatctg gggtccatt 180
tgaagaacca tttgtagttt tacagcttcg attctggaag agacaaacgt aacaaggagg 240
ttaagatac aaggattgaa atgtacggcc tgaagtgcag gggcatatga gtgtgggagg 300
tgcaagtggg gtttcttcta gaaaaactcc gatacaatag ggcatcaata tttctaggaa 360
gccacattct ccatagaagc tctcggttaag gggagctact ggtagtacag cagcatacag 420
ggggtgcagt gagagtgaag gggggtgaag gaacagtaaa aagaaaaata tgacaaggga 480
gggccaagag gatctacgat tctagttact ttcttcacgg ttgtcgctg aagagcaggc 540
gcagatcctc tagaggttca caggaatagc tagcattgtc tgctggattt tcgggttcc 600
ttggcagtat ccaggggttg gctcgagtgt gacttatcca agactccact ccagccactt 660
aactgcgggt agggtagata aaatgactgg gtaggggtcct tcccaggatg tgtgtaggga 720
tggggaatta aagggaagg gacttgacta ataccatgtc accagggtgg aataattcct 780
ttccctcctc tcagggaacag gttccctgta atgttttaag aactcgttga tatttggtta 840
aggaggtgat gctgcaact aagttggcgg tctctcagtc aagcacaagg tcattgggta 900
ggaagggtcg tccatacagc atctcatatg gactaagtcc tgctttttgg ggacagtttc 960
ggattcttag taaggctata ggcaacagag caggccatgc aagggtgggtt tcttgggtta 1020
gcttttttag atgtcgttt agtggttcat tcattttctc aacttttctc gaggatcgtg 1080
gcctccaggc acagtgttaag tgatattgta tacctaacgc ctgggatact ccctgcgtta 1140
ctgcagcctt gaaattgggg ccattgtcac tctgtaaac tcagggaagt ccgaatctgg 1200
gaattatttc atgaattagt acttttatta cctcttgggc cttttctgtc ctacaaggga 1260
aggcctccac ccaaccagtg aaagtaccca gattagtaga tactgaaatc tctgagattt 1320
gggcatgtgg gtaaaatcta gttgctagtc ttctcctggg taatggcctg ttctttgttc 1380
tcctgaagga gcttggcaat aaggcagggg attatttctt tggcacactt cacaggcctt 1440
gactatctgc ttgacagttt tgaaaaggcc tgggtccagta aataatgatt tggccatctg 1500
atgggtgctg tcaatgccta agtgaaaggc ctggtgaagg gttttaagta atttccattg 1560
gttagctgca ggcaaaagta ttttttctt ggtggctggc catcctgagg agaggaaact 1620
atgtcctcgt gaggttcccc attccatttc ttctgctgag tactggagct tggtttccca 1680
gaggggatta ccccatacta ggggtccttc tgtaagcatt tctaattggag agtcctgcct 1740

```

<210> 62

<211> 7140

<212> ADN

<213> Homo sapiens

<400> 62

```

ttggtcttaa gaacacaaat gatatggctc caatgactgg aggaacacca ggggtccttg 60
tctcacgctg atttagataa aacgactgtc aggcctctga gcccaagcta agccatcctc 120
ccctgtgacc tgcacgtata catccagatg gcctgaagta accaaagaat cacaaaagca 180
gtgaaaatgg cctgttctct ccttaactga tgacattcca ccattgtgat ttgttctctg 240
cccatcttaa ctgagcgatt aaccttgtga aattccttct cctgggtcaa aacctcccc 300
actgagcacc ttgtgacccc cgccccgtgc cctaagagaa aacccccctt gattataatt 360
ttccactacc cacccaaate ctataaaatg gccccacccc tatctccctt cgtgactcc 420
tttttcggac tcagcccggc tgcacccagg tgaaataaac agccttggtg ctcacacaaa 480
gcctgtttgg tggactctct tcacacggac gctcatgaca tttggtgcca aaacctggga 540
taggaggact ccttcaggag accagtcccc tgtccttgcc ctactctgt gaggacatcc 600
acctacaacc ttgggtcctc agaccaacca gcccaaggaa cagctcacca atttcaaate 660
aggtaagcag tcttttctact ctcttctcca gcctctcttg ctacccttca aactccctct 720
ctcaactacc tccaatctcc ctgtccttcc aattccagtt ctttttctc tctagtagag 780
acaaaggaga cacattttat ccattggacc aaaactccag caccagtcac ggacttggga 840
agacagtctt cccttgggtg ttaatcactg cggggacgcc tgctgatta ttcacccaca 900
ctccattggg gtctgatcac ggtggggaca cctgccttgg tcactacccc acattccctt 960
ggtggtacgt caactgcaaa agcaggggac gcctgcttgg gctgctcacc cacccttctc 1020
tctgtgtctc tacctttctc tttaaactta cctccttcac tatgggcaaa cttctgcctt 1080
ccattcccc ttcttctccc ttagcctgtg ttcttaaaaa cctaaaacct cttcaactca 1140

```

cacctgacct	aaaacetaaa	tgccttattt	tcttctgcaa	cactgctggt	ctgcagtaca	1200
aacttgataa	tagcttttaa	tggccagaat	atggcacttt	caattttctc	atcctacaag	1260
atctagataa	tttttggtga	aaaatggaaa	aatgggtctga	gatgcctgac	gtccaggcat	1320
tcttttacac	attgggtccct	ccctagttct	tgtctcccaat	gctgactcatc	ccaaatcttt	1380
cttctttctc	tcctgtctgt	tccttcagtc	tccaccccaa	gctctgagtc	ctttgaatcc	1440
tccttttgcta	cagacccatc	tgaactctcc	cctctctccc	aggtctgctcc	tcaccaggcc	1500
gagccagggtc	ccaattcttc	ctcagcctct	gctccccccac	cctataatcc	ttttatcacc	1560
tcctctctctc	acactcagtc	cggcttacag	ttctgttctg	tgactagccc	tcccccatct	1620
gcccaacaat	ttcctcttaa	agaggtggct	ggagctaaag	gcatagtcaa	ggttaatgct	1680
cctttttctt	tatctgacct	ctcccaaate	agtttagcgt	tacgctcttt	ttcatcaaat	1740
ataaaaaacc	agccagttca	tggcccatct	ggcaacaacc	cttacaggct	ttacagccct	1800
agaccctgaa	gggtcagaag	gccgtcttat	tctcaatatg	cattttatta	cccaatccgc	1860
tcccaacatt	aaataaagct	ccaaaaatta	aattctggcc	ctcaaacccc	acaacaggac	1920
ttaattaacc	tcacttcaag	gtgtacaaga	atagagtaga	ggcagccaag	tagcaacgta	1980
tttgagttgc	aattccttgc	ctcaactctg	agagaaaccc	cagccacatc	tccagcaaac	2040
aagaacttca	aaacacctga	actgcagcag	ccaggcgttc	ctccaggacc	acctccccca	2100
ggatcttgct	tcaagtgcgc	gaaatctgac	cattggggcca	aggaatgcct	gcagcccagg	2160
attcctccta	agccacgtcc	catttgtgca	ggaccccact	ggaaatcgga	ctgtccaaat	2220
cacccggcag	ccaatcccag	agcccctgga	actctggccc	aaggctctct	gactgactcc	2280
ttcccagatc	ttctcggtt	agcagctgaa	gactgacact	gcccgatcac	ttcagaagtc	2340
ccctggacca	tcacggatac	tgagcttcag	gtaactctca	cagtggaggc	taagtccatc	2400
ccctgtttaa	tcgatacagg	ggctaccac	tccacatcac	cttcttttca	agggcctgtt	2460
tccttttccc	ccataactgt	tgtgggtatt	gacggccaag	cttcaaaaacc	ccttaaaaact	2520
ccccctactc	ggtgccaact	tggacaacat	tcttttatgc	actctttttc	agttatcctc	2580
acctgccag	ttcccttatt	aggccgagac	attttaacca	aattatctgc	ttccccgact	2640
attcctgggc	tacagccaca	tctccttgcc	gcccttcttc	ccaacccaaa	gcctccttca	2700
tatcttcctc	tcatatcccc	ccaccttaac	ccacaagtat	gggacacctc	tactccctcc	2760
ctggcaaccg	atcacacgcc	cattactatc	ccattaaaac	ctaatacccc	ttacctgtct	2820
caatgccagt	atcccatacc	acaacaggct	ttaaagggat	tgaagcctgt	tgacacttgc	2880
ctgtctacagc	acgggcttct	aaaacctata	aactctccat	acaattcccc	cattttacct	2940
gtctaaaaaac	cagataagtc	ttacagggtta	gttcagaatc	tgcaccttat	caaccaaatt	3000
gttttgcccta	tcacaccctgt	agcaccacaac	tcgtacactc	ttttgtcctc	aatgccttcc	3060
cccacaactc	actattccgt	tcttgatctt	aaagatgctt	ttttcactat	ttccctgcac	3120
ccctcatccc	agcctctctt	tgtttttacc	tggactgacc	ctgacaccca	tcagtcccag	3180
cagcttacct	gggctgtact	gccgcaaggc	ttcagggaca	gccctcatta	ctcagccaa	3240
gctctttctc	atgatttact	ttctttccac	ctctctgctt	ctcaccttat	tcaatatatt	3300
gatgaccttc	tactttgtag	cccctccttt	aaatcttctc	aacaagacac	cctcctgtct	3360
cttcaacatt	tgttctccaa	aggatatcgg	gtatccccct	ccaaagctca	aatttcttct	3420
ccatctgtta	catacctcgg	cataattctt	catgaaaaca	catgtgctct	ccctgccaat	3480
tgcgtctcca	actgatctct	caaatcccaa	cctcttctac	aaaacaacaa	ctcctttccc	3540
tcctaggcat	ggttggtatc	ttttgccttt	ggatacctgg	ttttgccatc	ctaacaaaat	3600
cattatataa	actcacaaaa	ggaaacctag	ctgaccccat	agattctaaa	tcctttcccc	3660
actcctcttt	ccattccttg	aagacagctt	tagagactgc	tcccacacta	gctctccctg	3720
tctcatccca	acccttttca	ttacacacag	ccgaagtgca	gggctgtgca	gtcggaattc	3780
ttacacaagg	accgggacca	tgccctgtag	cctttttgtc	caaacaactt	gaccttactg	3840
tttttaggtc	gccatcatgt	ctccatgcgg	tagcttccgc	tgccctaata	cttttagagg	3900
ccctcaaaat	cacaaaactat	gctcaactca	ctctctacag	ctctcacaac	ttccaaaatc	3960
tattttcttt	ctcacacctg	acgcataatac	tttctgctcc	ccggctcctt	cagctgtatt	4020
cactctttgt	tgagtctccc	acaattacca	ttcttcttgg	cccagacttc	aatctggcct	4080
cccacattat	tctggatacc	acacctgacc	ctgatgattg	tatgtctctg	atctacctga	4140
cattcacccc	atttccccc	atttccctct	tttctgttcc	tcatgttgat	cacatttggt	4200
ttactgacgg	cagttccacc	aggcctgatc	gccactcacc	agcaaaggca	ggctatgcta	4260
tagaatcttc	cacatccatc	attgaggcta	ctgctctgcc	cccctccact	acctctcagc	4320
aagccgaact	gattgcctta	actcgggcct	tcactcttgc	aaagggacta	cacgtcaata	4380
tttatactga	ctctaaatat	gccttccata	tcttgacca	ccatgctgtt	atatgggctg	4440
aaagagggtt	cctcactacg	caagggtcct	ccatcattaa	tgcctcttta	ataaaaactc	4500
ttctcaaggc	tgttttactt	ccaaaggaag	ctggagtcac	acactgcaag	ggccacccaa	4560
aggcgtcaga	tcccattact	ctaggaaatg	cttatgctga	taaggtagct	aaagaagcac	4620

ctagcggttcc	aactttctgtc	cctcatggcc	agttttttctc	cttcccatca	gtcatttccca	4680
cctactcccc	cattgaaact	tccgcctatc	aatctctttct	cacacaaggc	aaatggtttct	4740
tagaccaagg	aaaatatctc	cttccagcct	cacaggccca	ttctattctg	tcatcatttc	4800
ataacctctt	ccatgtaggt	tacaagccac	tagtccacct	cttagaacct	ctcattttcct	4860
tccatcgtgg	aaacatatcc	tcaaggaaat	cactttctcag	tgttccatct	gctattcttac	4920
tacccctcag	ggattgttca	ggccccctcc	cctccctaca	catcaagctc	ggggatttgc	4980
ccctgcccag	gactggcaaa	ttgactttac	tcacatgccc	tgagtcagga	aactaaaaata	5040
cctcttggtc	tgggtagaca	ctgtcactgg	atgggtagag	gcctttccca	cagggtctga	5100
gaaggccact	gcagtcattt	cttcccttct	gtcagacata	attccttggg	ttggccttcc	5160
cacctctata	cagtccaata	acggagcagc	ctttattagt	caaatcacct	gagcagtttt	5220
tcaggctctt	ggatattcagt	ggaaccttcg	taccctttac	tgtcctcaat	cttcaggaaa	5280
ggtagaatgg	actaatggtc	ttttaaaaaa	acacccccacc	aaactcagcc	tccaacttaa	5340
aaaggaggat	agagcccaaa	aactcgcaac	caagctagta	attatgctga	accccccttg	5400
gcactctcta	attggatgtc	ttaggtcctc	ccaaatctta	gtcctttaat	atctgttttt	5460
ctccttctct	tattcggacc	ttgtgtcttc	cgtttagttt	ttcaattcat	acaaaaccgc	5520
atccaggcca	tcaccaatcg	ttctatacaa	taaatgctcc	ttctaacaac	cccacaatat	5580
cgccccctac	cacaaaatct	tccttcagtc	taatctctcc	cactctaggt	tccccgccc	5640
cccataatcc	ctctcgaagc	agccctgaga	aacatagccc	attatctctc	cataccaccc	5700
ccaaaatttt	tgctgcccc	acacttcaac	actatttttac	attatttttc	ttattaatat	5760
aagaagacag	caatgtcagg	cctctgagcc	caagccatca	tatccccgtg	gacctgcaca	5820
tatacatcca	gatggcctga	agtaactgaa	gaatcacaaa	agaagtgaag	atggcctgtt	5880
cctgccttaa	ccgatgacat	tccaccactg	tgatttgttc	ctgccccacc	ttaaactgagc	5940
aattaacctt	gggaaatttc	ttctcctggc	tcaaaacctc	ccccactgag	caccttgtga	6000
ccctgcccc	tccactaccc	acccaaatcc	tataaaatgg	ccccacccca	tctcccttag	6060
ctgactcctt	ttttggactc	agcccgcctg	cacccagggtg	aaataaacag	ccttggttgct	6120
cacacaaagc	ctgtttggtg	gactctcttc	acagggacgg	gggtgacaac	aacacggaca	6180
cacatggagt	ggttttaagg	agcagagagt	ttaatacgca	aaaaagaagg	aagaggctcc	6240
cctgtacaga	cacagagggg	gggggctcca	agccgagaga	aggaaacccc	atgtgcagtg	6300
gaaaagtggg	tgattatact	gggaggctgg	aggaggcggg	gtctgatttg	cacaggcccc	6360
aggggattgg	ggtgaccagg	tgtatcattc	atgtaccccg	caaaaaacct	ggccctccca	6420
cctcagccct	ttaatatgca	aatgtgggtt	gccatgatgt	tctgaaaaca	catgaattat	6480
ctggaggggg	ccatgacact	tggtacatgt	gctgacaaga	agaggggtgg	aatcgccatg	6540
gtggccatgt	tgggtggacc	tagtttttaa	tagcctgcat	ttgcatatca	aagtttgctg	6600
gcctggctct	tttaagctgtc	ttttctgtta	gaaaaggaat	ggtttggat	gggtgagggg	6660
tgcttcttat	tacaagaaaa	tttccaaaaa	cctttactct	ttctagctgc	caaaaaacta	6720
tttcttaata	acttatgtat	taccataatt	aggcagcacc	aaagatccct	gcaggtcaga	6780
ccactgcaat	taacatgctg	gctttactgc	tgattatggg	agctgcatcc	acctagcctc	6840
tcatattgca	actgcctgac	ctctgccacc	ccacgagcca	cttatcccca	cttataatca	6900
gcccatttgc	attgtaacat	ctgccactta	ttcccagcgt	tgtggtatat	cctatagatg	6960
aattcattca	acatccattc	caacaccacc	tctcttgctc	tcctatactc	tctggagagt	7020
gaattactga	gtcacatgat	cttcaactgca	gtcatttgtg	gctatgtgac	atagttctgg	7080
acagtgaaca	tagacagaag	tccttggggc	gggcttctct	tctgggatga	gggcaaaacg	7140

<210> 63

<211> 44100

<212> ADN

<213> Homo sapiens

<400> 63

tgccttttatt	tccgtaggct	ggtcatatgg	cgctagcact	cacataaagc	taccgaggag	60
agcgaatgaa	accaaaaatca	ctttaccttc	acagcacgag	gccgtcgtcc	ctctcgatat	120
ttggcccgtg	tgtcgcatac	cgccctctgg	acgtgggtgat	caaataaaact	ccctagctcc	180
ccgcccgtcg	acgccatctt	gcctactttg	atcctcgcag	ggaggacaac	atccgcccta	240
ctgagctccc	ttttatccaa	taagagagcg	ggatgagtta	aggagtgcc	ggattggctg	300
gagaatcgac	agcgtcggcc	atcgtttcct	gcgtgcgaag	atttgatgaa	cgagggtgcc	360
cccccgagcg	gctcggcgga	gaggcgcggt	gggtgacaga	agctttcttg	tcccaccac	420
tacaggctta	cggcaggatg	cgcagcgggg	agagggggcg	gggccgcagg	gggcggggcc	480

gatcgatctc	ctccggctcc	gacgtcctcg	gcctgccggg	tcccgggtcc	tttgccggcg	540
taggggtggg	gaacccagag	cgacgctccg	ggacgatgtg	gggcagcgat	cgccctggcg	600
gtgctggggg	agggcgggcg	gcagtgactg	tggccttcac	caacgctcgc	gactgcttcc	660
tccacctgcc	gcggcgtctc	gtggcccagc	tgcattctgt	gcaggtaacc	tgcgggcccc	720
gagccacctg	atcttcagcc	tggggtcgga	cgaggccgaa	gcctctcagg	gacggcgccg	780
gacaccggct	gccacccggg	cgccgccgaa	gcgcgcagag	atcaggggtc	ctcgacggca	840
gggcccctct	gggtagtctc	tggatcccac	aagtcacgtg	cagccctggg	ctcgtcttat	900
cccaggctct	ttcacttggt	gaaactgaac	ctagaaacgt	cctaatatcc	taccactggt	960
tttataaata	ttccttattc	caggctggaa	aagctcctga	gaagtgggtt	gtttttatta	1020
ttttaaaagg	tgttttcctt	gccagccatt	tccagttaac	ctgcgctgct	gccgtccggg	1080
ccgcgagagc	gggacgcaga	gttggtggcg	gagccctgt	cggttcccgg	ggactaagca	1140
ccgcgtccca	tgagcgggaa	aggttaatac	aatgatgggt	ctgccctgcg	tcgctgacgc	1200
ggaacacagc	tgtagtgtgt	taggaacaca	taacgtagtt	aagatcactt	gaagctctgc	1260
gatcagtcgc	ccttctggac	gttggtggtt	ggatgtttca	cagttctaac	cactgggtga	1320
gatacagcgt	ccataatttt	ataattaaaa	atagaggcac	atgggtctac	gagtttgagt	1380
gtacttatgg	gggcaaaaagg	acggcgtatt	tgaatcctc	ataaatcctg	gatgcattgt	1440
accaccaggt	ggctaatact	tgcaatgaat	agagtttgca	ataatttcaa	gcatcccttc	1500
tttccacttg	agttactttc	ccatacctta	gggaagatat	ttttggtcca	ctgaaaacat	1560
gagttcagca	gaatcctcct	atcatcgtcg	ttattttttt	ttaccactaa	gtagacaatc	1620
ttttggtttt	tgatgggctt	tatggctaga	gacaaatcag	tcactgtcac	caagttccag	1680
gtagaagttg	gttcagtgct	ctgtcagctt	cgatgggatt	tttcaacatg	ttttcaaata	1740
tgcacttaat	agtaggaatg	ctttcttaca	gtaactctaa	tttgatccca	agatgtagtt	1800
gttaccttac	attcatcact	gtttaagaat	ttagtggctt	tgatctttgt	tttaaatttt	1860
gagccttcgg	gaagtactta	taagaattaa	ttcatgcata	tctttttgaa	atgtaaatgt	1920
ccttagccct	ggaacaaaatt	gctgtttctg	ttcagcccat	attagcagaa	taggtcaact	1980
ttactttcta	attatcaatg	taataagttt	attactttat	agattccata	aatctataca	2040
tttattcctc	gatgaattat	ataaattttat	agaattttatg	ttttatagaa	aatttggaat	2100
gcatggaaaa	ttattaacaa	gaaaataagt	taccataat	cccagaactt	agaggtgact	2160
aatgttgaca	gtttggatca	aatcttccag	ttttgtttct	aatctttatt	tttaacataa	2220
atgaggtcct	gtatacacac	gtacagtttt	gtgtcctggg	gtttttattt	aatgttatta	2280
tgaggtgttt	attttggtta	aaggtcatca	ttttaagttg	tttaattagta	ttctagcaca	2340
aatttgccat	aatttatatta	attgtttact	atgattgacc	atttagattg	tacttaattt	2400
ttaggcatta	gaagtgataa	actatatttt	aatcagacgt	tgaataaac	acatctttgt	2460
ttagaaaaca	tcattttatt	tctgggtgtc	taggatagat	tcccagaatt	cttgggttag	2520
aggccataga	taattatgaa	agcagaaaga	ttcacaagtt	gggagttaat	acttgaatta	2580
ctttatttgg	ggtgaagcat	tgagtgcata	atacagatca	tgagtaatg	ggaagaagg	2640
ttggaacaat	ggttttctgg	cctatgtcag	acttaccttg	aagcttttaa	gaatacagat	2700
gttctgatca	accctcagac	ctattaaatc	agacctaaaa	tcttagggaa	taggcttttag	2760
gcatctctaa	ttttaaaaaa	tttattcagg	ctacttggtg	gcacaaaaga	gttgagacct	2820
actgtcctag	aatcatagaa	ttttaatgac	gatagagacc	ttaagcatct	aggtcggttc	2880
tgtactttta	catgtaagga	aactggcatt	cctaggccag	taccattgcc	atgcagctaa	2940
tttgccctct	tgtctatagc	tcactctgca	tcacccaacc	taccgttctc	actgtttctt	3000
ctataacca	tctccttccc	acttctgttc	tcttactcat	gccattcttc	cctcagtcac	3060
ttttcttctc	tccatacaaa	ttccatgtct	ttaaaaagga	ataatcctac	ctcctccaca	3120
tagctttcca	attctctgtt	gcccacattt	gtctcccttt	caatacttct	ctgttggtgt	3180
atgtgacaca	tcacatttga	tatactctgt	actgtgtttc	aagtattgta	ttctcttggt	3240
tactcaagtc	attatttcag	gactgactac	ccagtatgat	ctttaagtca	ggatttctca	3300
accttgggac	tggtgacatt	ttgagctgga	taattttttg	ttttgggggc	tctcctgtac	3360
attttaagat	gtttaacagc	acccttggcc	tctatccagt	agacgcctgt	actgcctccc	3420
cctatctgtg	acaacaaaaa	aggtcttcag	acattgtcag	atgtctactg	aaggacaaaa	3480
tcacctctgg	ttgagaacca	ccgcttcaac	taagttatct	tctctgtact	cagaacttga	3540
tgtgattgca	gcagggggag	aggattcata	tacacagtga	atgcaaacga	acctaataca	3600
ccattcgcat	atggccacac	aattttccatt	tcccttggtg	tagcaagaga	taccctaggc	3660
tttggaacctg	attattccta	aggcattctg	atgtatgggt	ttacctgcag	atttccctgt	3720
aatactgata	cctcagtttg	ggtcaaagaa	ggtcaattaa	ttgattgatt	tgatttgact	3780
cctggaaaag	acgctccttt	ctagctgtct	ctttcttctc	tttacctgaa	tagccagggc	3840
tctgtgggtc	aagtgaagta	ttttgacata	aaaattaact	tagaacattg	gtctgcagag	3900
tttgctcaat	ataactgagc	acatattgtg	gctttatgga	gctggttact	actttttgac	3960

caaataaata	attagaagta	tttttcctcc	tcaataaggt	tcatttttcc	ttttttcagt	4020
gagctggtag	agtttccttt	tttgatat	cagggcatct	ttcataattc	catctcttaa	4080
gtttcttcat	atgaagtaga	atztatctgg	attatgtatt	gctgactctg	atgaaaaccc	4140
atagaaaagca	tctggggcct	gatcaccttc	attcttgtaa	tagctcacac	ggttacagct	4200
gatatggtaa	cttaagactt	ttgattccaa	atctaggcaa	aatacactca	gttgaaagaa	4260
tttgtcagcc	agaacagttg	gactgttctg	tgaaaattgt	gagaaaaatt	acacaactaa	4320
gtgatacatg	atgatggctt	tcttaaatat	aaaattgtaa	taacatgggt	aattttccagt	4380
acgttatatt	gtcccagaag	tggctccaac	attgtttgaa	atttgtctca	tttaaagaaa	4440
cataagctgg	ctatgggtggc	tcacgcctgt	aatcccagca	ctttgggagg	ctgaggcagg	4500
cagatcacct	gaggtcagga	gttcgagacc	agcctggcca	acatggtaaa	accccatctc	4560
tactaaaaat	acaaaaatta	gccgggcatt	tgggtggggc	ctgtaatccc	agctacttgg	4620
gaggctgagg	caggagaatt	gcttgaatct	gggaggtgga	ggttgcagtg	agccgagatt	4680
gtgccactgc	cctccagcct	gggtgacaga	gtgagctctc	gtctcaagaa	aaaaaaaaaa	4740
aaaagcaaga	aacataaaga	ctgggcatgt	tggctcatgc	ctgtaatccc	agcactttga	4800
gagactgagg	tgggaagatc	acttgagccc	aggaggttaa	ggctgcagtg	agccgtgatt	4860
ttgccactgt	actcgagcct	gggcaacaca	gtgagatcct	gtctcaggaa	aaaaaaaaat	4920
gcatgtaaat	gaatgaattt	gatattttaat	attttaaatt	atgaaaactg	ttctgtagag	4980
atgtagatct	tgccatgttg	cccaggctgg	ctttgaactt	ctgggctcaa	acaatcctcc	5040
tgtctcagtc	tcccaaagta	taaagattac	acatgtgagc	cactgcacct	ggcctaatat	5100
ttttaactta	atgaatttat	tttgatataa	ataaattaat	aacactgaag	cttcctgata	5160
taataagtct	ttttgtgtgt	gtgacgggtt	ctcactctgt	tgcccagact	ggagtgtaat	5220
ggcactatca	tggctcactg	tagcctcaac	ctccctgact	caagtgatcc	tcccacctcg	5280
gcttctctgag	tagatgggac	cacaggcgta	tgccaccaca	cctggctgat	ttttaaaaat	5340
tattattgat	acatattaat	aaaattat	ttatttttaa	aatgatatat	gtggctgggc	5400
atgggtggctc	atgcctgtaa	tcccgacagt	ttgggaggcc	gaggtgggag	gatcacttga	5460
gaccaggagc	ttaagaccag	cctaagcaac	atagttagat	cccatctcta	tagaaaaaaa	5520
aaatggctag	gtgtgggtgt	gtatgcctat	attcccagct	actcaggaga	ctgaggtgag	5580
aggattgcta	gagcccagga	gtttcaagtt	acagtgcact	atgattgtgc	cagtgcactc	5640
cagcctgggc	aacagagcaa	aatcctgtct	caaaaaaaaa	aaaagttcga	aaatgcttat	5700
gatgcaatat	aagtagtgga	aaaggatat	aaattgtgcc	tatatgaaca	caactatatg	5760
aaaaacttgc	acatagagaa	aaggattaac	aagaaataga	ccaaattgtt	cacatggttg	5820
tcttgtttgt	ggagagaata	tcagtagttc	atttgtttcc	ttccaagttt	atatgttttc	5880
cgaggtctct	ataatgagtt	tgtaattgtt	taatcataga	aaaccctttt	ttggctcctg	5940
gccacaaact	tacatgtttt	aatgtaattg	cttttttaat	gagaataaat	gttatatttt	6000
gcttttttaa	aacctatatt	cccatagtta	tatgagccct	tacaattatt	aagaggctgc	6060
ataatataac	ttttctggaa	gggtacagaa	gaaacagcag	taattacctc	tgagaacaga	6120
gacatggcct	cacattttac	ccttttgtac	agtttgtgct	tttgccacat	gcattttatta	6180
ttcttccaat	aaataagtaa	ataaatatgg	attgtatact	ccatctgggt	gggtgtttcat	6240
aattctaaaa	ttatatgtct	acatttttaa	agatgatatg	tgtttctact	tattaacgta	6300
tatgttaaaa	tagtaaat	atatcttatt	taataatttc	cctattgata	gacatttaag	6360
acagtctcaa	gtgttcacta	tcatagaaaa	tactgcacag	atagcttttg	ctatagtttc	6420
ttttttcttt	gaatcgttaa	ttgggaataa	atgctcaaat	agttatatgt	ggctcaactg	6480
ctattttaagt	ttattgactg	actgctgcca	ttttgaattc	tgaaggggtt	gattaaaatt	6540
ataatgctgc	cataagaata	taagggtatt	ggcttcatta	gcatccacca	gcattgggtg	6600
ttggaaatga	ttatagattt	ttaaatgcta	caacaaatgt	agataacaga	gaactatcta	6660
tagaactctt	tttgacatg	tgaattgtaa	taatagttta	ttttcatgtg	aatccagaaa	6720
aatgtatacg	aaaacctttt	ttcctctcat	ttcttatatg	aatagaatca	agctatagaa	6780
gtggtctgga	gtcaccagcc	tgcatctctg	agctgggtgg	aaggcaggca	ttttagtgat	6840
gggggacagg	taagcacatg	tgatggcaat	aactttcttc	taatatacaca	taatatagca	6900
atagaaataa	aattaaaagt	ttagattttt	tgtaaaggga	ggtgagatgt	cacctaat	6960
gtatgctatt	atgtaactag	tctaggatat	tgaagctgac	tatactctgt	ttttaggtca	7020
ttatcttgta	gtttaccata	ctccctactt	gcttcttatt	ctactattta	actcattttc	7080
cacatccctt	aattttgggt	tcatgaaatt	atttttcctt	ctgaattact	aggttctact	7140
tactattatt	aaactttatt	tctgacatat	tttataacct	tccatggctc	cacttgatta	7200
aaaataaaaa	attcagctgg	gtgcggtggc	tcacacctat	aatcccagca	ctttgggagg	7260
ccaaggtggg	cggataattt	gaggtcagga	gttgagacc	agcctgcca	acgtggtgaa	7320
acccccctc	tctactaaaa	attcaaaaat	tagctgggca	tgggtggcagg	tgctgtaat	7380
cccagctact	caggaggctg	aggcaggaga	attgcttgaa	cctgggaggt	ggaggttgca	7440

gtgagctgag	attgcactgc	tgcacttcag	ctgggtgaca	agagcgaaac	aatgtcttga	7500
aaaaaaataa	aaaataaaaa	attctacaac	acaggggttat	tattttttcca	tttttgtttt	7560
cccttatgag	tttaatatgt	ttagattata	aacctgaaag	cttgaatacc	tatgtctatc	7620
ttttgttttc	ttatgtttat	caagttattc	ctttaaacat	tttctaaact	gtaagaataa	7680
tgtgaggctg	ggctcaatgg	cttatgcctg	taatcccagt	gctttgggag	gccaaaggtg	7740
gaggaccact	tgaggccacg	agttcaagat	tagcctggct	aggcaacata	gcaagaccct	7800
atctctataa	aaaaattaaa	aaaattagct	gggcatggta	gcaaatgctt	gtagtcaccg	7860
ctactcagca	gactgaggta	ggaggaatgc	ttgagaccag	gaatttgagt	gacctatgat	7920
tatgcactcc	agcccgggca	atagcaagac	cctatctctt	aaaagaagaa	gatgtagtaa	7980
taatacatat	tcattataac	tattttacca	ttgaaagtaa	aaaatgagtt	tttacctttt	8040
cccagtccca	tcctcagaat	ggggatctca	gtagaccttt	aggattggaa	gaatgagatc	8100
attcataatt	tctgcaatta	ttaccccaca	aaatatttca	gatacctttc	catgtattac	8160
aaacaatgtg	catttaacat	gtctctctct	ttctctctct	ctctgtgtgc	gtcttcatga	8220
tcctctgttg	cagccctgcc	agtaagacac	tatctcctga	agaatcactg	ataggaaccag	8280
aaagtggact	ggctaggcca	ggagtcctta	gcttcttagg	gggcaggagc	tgctttgtgc	8340
tttctcagaa	tcagatatat	atgtggactg	aaacatttaa	aaacagaata	gccaaaggtg	8400
ctatacgttt	aaaacttata	tagatggggc	tacattgtct	tctattacta	atttcccatg	8460
acaatacacg	agagtgccat	gtctttttta	cttgttttga	gcacagacta	atcttgttta	8520
tgcatgtttt	ttgatgagaa	taggctactc	atgagaaatc	tgtaaaccta	acactagtcc	8580
cttgcatact	ctaaattgtt	gctagaatct	taaaatttta	gcaccagacg	gaccttagaa	8640
atcattaact	ttggtgcttt	gttctacaat	acaaggagat	ggaatatttt	accaggattt	8700
gcttagcagg	ttacagttct	gccctctgag	taccagcac	ttccctgtgg	gcaacatcaa	8760
cttctcgatt	ttcaagtctt	aattagtact	ctgaagaatc	ctacttgttt	ttaactccca	8820
tttgctttga	agtgacttta	cctgattttt	ttagatccct	tattgcagca	atgccactaa	8880
gaaactgagt	ctctagcttc	ttggtgggca	ggagctgctt	tgtgcttgct	cagaatcatc	8940
cttttcagta	agggagatat	tgaagagaaa	tctactgagg	agtctggggg	tgaggcactc	9000
agggaaatcc	tgctccagtc	cacaaaagca	gagaggaagg	gttgggttacc	tagagtattt	9060
aacatgcaga	ggctttggat	tttactcctt	taatccttgg	aaatgcctat	ggaaggggaa	9120
aggaagtaag	atggtgactc	cagcttatag	acatactagt	gttacaatata	tttaacttat	9180
aatgagagg	tatttattagt	tttacttaac	tttcaactgt	gaaggattat	acttctcaat	9240
atttgtctcc	agtgtctatt	tcagtgtatt	tctcactttt	cttgaagcag	actgtgtctt	9300
gcaaaacttc	tagaaataat	gagaatattt	atatattaga	tcaagccata	acttgatgat	9360
atagtcattt	cttcttatat	tttttactta	cattttttaca	ttttaatgat	tactttcatt	9420
tttgaaaaac	atgtcatgct	gagatgtatt	tttcttcatt	ctgtaattag	ttatgaaaca	9480
gtttttccta	aaatgctgag	tatatcaagt	cttggctaag	aataagtaat	aaatatttgc	9540
cacatgaaag	actacacata	tagccaggtg	cagtggcttg	cacctgtttt	cccagctacc	9600
caggaggctg	aggcaggagg	attgcttgag	cccagggttt	ccaggctgca	gtgaactatg	9660
attgtaccac	tctactccag	aatgggtgac	agagccaggc	cccatctctc	aaaacagaaa	9720
agaaagatta	catagactac	atatacaccc	ccatccaaaa	catacacaca	catctactta	9780
acctaaaatg	gtaagaagat	aacttcttat	tttctaatat	atgacacaga	aaagtttttt	9840
taaagtagtt	ttaaattttt	aattttttct	aggatattct	caagccatgt	tcccatgtgg	9900
tatcttgtca	acaagttgag	gtggaacccc	tctcagcaga	tgattgggag	atactggtaa	9960
agaaaaccaa	ataagaacta	tctcatttaa	ggttaaattta	cttcacaata	tcaatgtctt	10020
tagctttctc	taagctttat	tatatattct	gagttggttt	tgaattataa	gaatgaattg	10080
gggccaggca	cagtagctca	tgccatatagt	cccagcactt	tgaggaggcca	aggcaggttg	10140
attgcttgag	tccaggagtt	caagaccagg	ctgggcaaca	tggtgaaacc	ccgtatctac	10200
taaaaaataa	aaaattagcc	aggcatggta	gtgcatgcca	ttagtccag	tcacttggga	10260
ggctgaggca	ggagaatcgc	ttgagcccg	aaagtcaagg	ctgcagtgag	tcaggatctt	10320
gccattgtac	tccagtctgg	aaaacagagt	gagacctgt	ctcaaatata	aaaagaatga	10380
attgatagag	atctaagtga	caacctgaca	actataggta	ataaaattgt	attggggatt	10440
catgtttaat	gagtagattt	taactactct	taccacaaaa	acacaaaagt	gggtaactgt	10500
gagatgatgt	atatgttaat	ttacttcact	atagtaacca	ttatactatc	tatatgtagc	10560
tcataacacc	atgtcgtgta	tattaaatat	gcacattaaa	atttgttttt	taaaaaaaga	10620
attgagattt	tttttaacta	gatattggagt	ggacaaaatg	taaaagtgaat	tgatcttttc	10680
gtctgttgg	tctaggagct	gcatgtgttt	tccctgaac	aacatcttct	agatcaaat	10740
cgaatagttt	ttccaaaagc	catttttctt	gtttgggttg	atcaacaaac	gtacatattt	10800
atccaaattg	gtaggtgcta	ttgtaatat	tgctgtcata	ttctacacta	tagcattgag	10860
tccaaagtag	aaatgaatgt	gcactaatga	gctttatttt	ctacacagtt	gcactaatat	10920

cagctgcctc	ttatggaagg	ctggaaaactg	acaccaaact	ccttattcag	ccaaagacac	10980
gccgagccaa	agagaataca	ttttcaaaaag	ctgatgctga	atataaaaaa	cttcatagtt	11040
atggaagaga	ccagaaagga	atgatgaaag	aacttcaaac	caagcaactt	cagtcaaata	11100
ctgtgggaat	cactgaatct	aatgaaaacg	agtcagagat	tccagttgac	tcatcatcag	11160
tagcaagttt	atggactatg	ataggaagca	ttttttcctt	tcaatctgag	aagaaacaag	11220
agacatcttg	gggtttaact	gaaatcaatg	cattcaaaaa	tatgcagtca	aagggtgttc	11280
ctctagacaa	tattttcaga	gtatgcaaat	ctcaacctcc	tagtatatat	aacgcgtcag	11340
caacctctgt	ttttcataaa	cactgtgcca	ttcatgtatt	tccatgggac	caggaatatt	11400
ttgatgtaga	gcccagcttt	actgtgacat	atggaaagct	agttaagcta	ctttctccaa	11460
agcaacagca	aagtaaaaaca	aaacaaaatg	tgttatcacc	tgaaaaagag	aagcagatgt	11520
cagagccact	agatcaaaaa	aaaattaggt	cagatcataa	tgaagaagat	gagaaggcct	11580
gtgtgctaca	agtagtctgg	aatggacttg	aagaattgaa	caatgccatc	aaatatacca	11640
aaaattgtaga	agttctccat	cttgggaaaag	tctgggttag	tataaatttt	ataacttggg	11700
agaaatttta	tgtggcttaa	acatccccaa	attatgaatt	agaatagtat	ttcatatata	11760
aattgaaaat	caattaaaaa	gaaacacagt	gcctaaaggc	acttggggga	cacattttacg	11820
ctttgcagta	aagtccttgt	ttggataaag	attgtatggt	ttctggccaa	gtaagcttga	11880
ataggtacaa	gcttagatag	gttcaggcca	gagaggtaa	aattacttgc	ctgagattgc	11940
ataggtcttg	ttacaactag	gattcaaact	caggcagatt	gacttggggg	ttcagcagga	12000
tggagtggcc	tacaaagcct	cccatcttta	atgcttgcat	atttgttccc	cagttaccga	12060
aagcaacttg	ttaatatatt	ggaaaagggc	cagtgtaggg	agagatccat	ggcatgaggt	12120
aaccttcctg	ctgcatgtgg	tggcacctgg	attggaatgc	atccaggagc	tgcttaccct	12180
gccggtgtct	gctctttaat	ttgtgtataa	cggagaggaa	gtagacaggg	caactagtgc	12240
tccagccctt	catcctggcc	acaaatatta	atgctacctt	tatatgacat	aagtcactag	12300
tccattttatt	ggaacctaaa	tttgaaccac	tgtaaagtaa	gacttcatag	tgataaagag	12360
aggaacttgt	taggaaagag	aataaaatag	aaagagaagg	ttgtctcctt	ttgtagattt	12420
tttttttttc	tccaacagtt	ttacctgtga	cctttataca	aataactgac	aaagcattaa	12480
tctctttggc	ctacatcatt	ttcttttcta	tttttttttt	ccacaagatg	gagtttctact	12540
cttcttgccc	aagctggagt	gcagtggcat	gatctggctc	actgcaacct	ccgcctccca	12600
cgttcaagtg	gttctcctgc	ctcagcctcc	tgagttagctg	ggactacagg	catgcaccac	12660
cacgcctagg	taattttttg	tatttttaagt	agaaaactgg	tttcaccatg	ttagccagcc	12720
tgggtctggaa	ctcctgacct	caggtgatct	gcctgcctcg	gcctcccaaa	gtgctgggat	12780
tacaggcatg	agccactgct	cctggccggc	ctacatcatt	ttctaaagct	ccagaccatt	12840
cttttctttt	cttttctttt	cttttctttt	cttttctttt	cttttctttt	cttttttctc	12900
ttctctttct	ttctctttct	ttctctttct	ttctctttct	ttttcttttc	tttttttgag	12960
ttagaagctt	gctttgttgc	ccaggctgga	gtgcagtggc	accacctcca	ctcactacaa	13020
cctccacctc	ccaggttcaa	atgattctcc	tgccctcagc	ttcagagtag	ctgggactac	13080
aagtgtgcgc	caccactcct	ggctaatttt	tgtattttta	gtagggacga	ggtttcacca	13140
tgttggccag	gctagtcttg	aactcctggg	ctcaagtgat	ccgcctgcct	cagtcctcca	13200
aggtgctggg	attacaggcg	tgagccactg	tgcttggcct	cagatcatta	ttttctgtta	13260
gctttaaaact	gtccgttcag	gagatcccac	tgcatcctca	aattcaaaat	atctaacaact	13320
gagcttatga	tttagctggg	tctgtcatta	gatgggaata	tcctttttat	tccttgaaat	13380
tatatggtga	gaacagggag	aagtgcctgat	ggtaaaagtc	tgtgatttaag	atagcaataa	13440
ggactccgcc	cttcccactc	cactgaagggt	tgaagagcca	tggaacaatga	gaagtcacag	13500
taggtgaaat	caggtactaa	aatggacttg	gcttgagaga	tcaaaattga	tcacttggtg	13560
atacaactaa	caaattcatg	ttaacttgaa	cctttattac	cctgtgaagc	atggtgatta	13620
aaaaaaaaaca	acaaacaaac	aggaaaacttg	attgttaaatt	tctctttaag	tcagaatatg	13680
taccttagag	tttttattta	tgctttttgtc	taccattaat	atgtctgcac	ctgctcttta	13740
gaagttaata	gagagtaaag	tcgtcttttat	gtctttcagt	gcttacttat	atttgggaag	13800
ttgagaaaaa	tttttaacat	cattattgat	atatatatat	atatatatat	atatatatat	13860
atatatatat	atatatatat	atagataatt	tttttttttt	tcttgagacg	gagctctact	13920
ctgtcgccca	ggccggagtg	tgggtggcgt	ctccactcaa	tgcaagctct	gcctcccagg	13980
ttcaagcgat	tctcttgcc	cagcctcccc	agtagctagg	atacaggctc	ccaccaccac	14040
gccttgctaa	tttttgtagt	tttagtagag	acgaggtttc	accatatttg	ccacgctggg	14100
ctcaaaactcc	tgaccttggt	atccgcccac	ctcggcctcc	caaagtgtcg	ggattacagg	14160
cgtgagccac	tgcgcccggc	tgaggtaaaa	tttaaagtgt	acaattcagt	catttttagt	14220
atattttatac	tagttgtaca	gccatcacca	caatctaagt	ttagaacatt	ttcattaggg	14280
ggtgggagaa	attttactct	gcttttttaga	ttaagtttct	gtctggatct	aatcatttaa	14340
tcagacaatc	aggcagattg	tctgtgatta	gttttgccca	ttccagcttc	ttcattgggt	14400

gttaactttc	acaaataaaag	gctgctcaaa	gattagaaat	aacattttaat	ttgaatgtaa	14460
atgtgccata	gtttaaaaga	tgggtttggt	gaatacagtc	aaatacacatc	atttaaagct	14520
ctaattctga	agattatgta	aagaaaagga	aagaaatgta	gggagaggat	tgaatgttc	14580
atggtataac	aatatctgaa	catccatctg	gtcacaccgt	tggatattga	atgttttgtc	14640
ctcctcaaat	tcatatgtcg	aaatcccaac	tcccaagggtg	atcgtattag	gagggtgtggt	14700
ctttgggaag	tgattagggtc	atgaagggtga	agccttcacg	aatgggattc	gtgctcttat	14760
aaaagagaac	tgtgagaaat	aagtttctgt	cgtttggttag	ccaccaggtt	taggatattt	14820
tgatatagca	gcctgcatgg	actgagacaa	ctatgagtta	ttatgatagc	ttctgttatt	14880
tcacctaaat	tcatagaagc	taatatatca	atatttatgc	tatgaaatat	ttcttaacca	14940
agctttgaat	atatttatat	ttttgtttat	ttttaaattt	cagattccag	atgacctgag	15000
gaagagacta	aatatagaaa	tgcatgccgt	agtcaggata	actccagtg	aagttacccc	15060
taaaattcca	agatctctaa	agttacaacc	tagagagaaat	ttagtgaagt	caaatatata	15120
tgttacatca	aaattctttt	acacgttttg	taagatttct	agttgcttta	gctaagtaat	15180
aagaatgttg	tattcctttt	tgatacaaat	ctttttttat	tgtgttaaac	tatatataac	15240
ataaaatatg	ccatgttcgc	catttttaag	tgtataattc	aaaggcatta	attacattca	15300
taatattgta	caaccatcac	cactatctat	atccagaact	tttccatcac	cccaaagaga	15360
aacttggtac	ccattaaaca	ataattcccc	gtccactcct	ttccccagtc	cctggtaatc	15420
tactaatgta	atgtgtctc	tatgaattta	cttattctag	atatttcata	tataagtagc	15480
agtatgcatt	tgtcttatgt	atctgactta	tttcatctta	cataatgttt	tcaaggctca	15540
tctgtgttgt	atgtatcaga	atgttattcc	ttttcatggc	tgaatactat	tccattgact	15600
gcataataca	catttggtta	tccattcatc	tgttgatgga	cacttggtgt	gtttccacat	15660
ttttggctgc	tgtgaataat	gctacagtga	acattgggtg	acaagtatct	gtttgagttc	15720
ctcttttcag	ctcctttggg	atatacctag	gaattatgtt	taactttttg	agaagctgag	15780
aaatctttaa	taaatgataa	cacaaatact	tatatattgcc	aatgcaaata	tgaatatattt	15840
tggcttttaa	gagattgata	attttgccac	gtggtgttaa	ttaaaaaaaa	ttgtcccatg	15900
ttgtttcagt	attaatattg	tagcctaaaa	gagtgctaga	ctgttttact	ttttactcag	15960
ttaattcttt	ggatactggt	agagtcagga	aatgagatat	tgaacttaaa	gatctttgca	16020
ggtgggggtcc	agtgggtcac	acctgtaatc	ctagcacttt	gggaagctga	ggtgggagga	16080
ttgcttgagg	ccaagagttt	gagaatagcc	tgggcaacat	agcaagaccc	catctctaca	16140
aaaaaattaa	aaaaaaaaatt	aagccaggcg	tggtagctca	cgctgtttat	cccaacactt	16200
cgggaggctg	agatgggtgg	atcacttgag	gtcaggagtt	gggagaccagc	ctggccaaca	16260
tggtgaaacc	ccatctctac	taaaaatacc	aaaattatcg	gggcgtggtg	ctaactctgt	16320
aatctcagct	actcaggagg	ctgaggcagg	agaaccactt	gaactgagga	ggtggaagtt	16380
gcagtgaagc	tagatctcac	cactgcactc	cagcctgggt	aacagagcga	gactctattt	16440
caaaaaaagt	aaaaataaaa	attagacaca	tgtggtggca	catgcctgta	gtcctagcta	16500
ctcaggaggc	tgactgaagt	gggagactct	cttgagccca	ggagttccac	actgcagtga	16560
gctatgattg	tgccactgca	ctccagccta	ggcaatatct	caaaaaaaat	ttttttaaat	16620
agattattag	gccagacgtg	gtggctcatg	ccagtaatcc	cagcactttg	gaaggccaag	16680
gcaggcggat	cacctgaggc	caggagtttg	agaccagcct	ggccaacatg	gtgaaacccc	16740
atgtctacca	aaaatacaaa	aattagctgc	aatgtctata	atcccagcta	cttgggagcc	16800
tgaggcaagc	gaatcgcttg	aaccggggag	gcagagggtg	cagtgaagtg	agactgcgcc	16860
actgcactcc	agcctgggag	atacagcgag	attctgtctc	aaagaaaaag	gaatttgttt	16920
tcctgtcttt	atcgtagagg	gaggaaaggg	agaatggggt	tggaatgggt	attgagttag	16980
ccacattatg	gtagatgtat	cactgggcat	agagaaaagg	agcatttaaa	actttttccg	17040
ctaacagatg	tttcttcagg	ctacactgca	ctcattgtgc	taactgtaat	gtcaaatccc	17100
agacctgtgc	ctatagaaca	tgaacatcct	tcattggatt	tgtttggtca	ggcttacact	17160
ttattaggaa	gatcagatgt	taaaataagg	gtgttaaagt	taagttcaga	tatgaggata	17220
attcattact	attccttttt	ctggcagcct	aaagacataa	gtgaagaaga	cataaaaaact	17280
gtattttatt	catggctaca	gcagtctact	accaccatgc	ttcctttggt	aatatcagag	17340
gaagaattta	ttaagctgga	aactaaagat	ggtgagtaca	tttgttattt	tgactttttt	17400
ttctatttaa	atagttgtac	atttttaatt	gttcttgcaa	cctgtcatatc	ctgtgaaacag	17460
tatgtgaata	gtgaaatata	attatgataa	ttaaacagta	gttttttatgt	attgaaaaat	17520
atctttggcc	gggtgcagtg	gctcatgcct	gtaactccag	cactttggga	ggccgaggca	17580
ggcggatcac	ttgagccag	gagttcgaga	gcagcctgcc	aacatggcgc	aacctatct	17640
atacaaaaaa	atacaaaaaat	tagcctgaca	tagtggtgta	tgctgtagt	cccagctact	17700
tgggaggctg	aggcagaagg	atcacttgag	cccaggaggt	ctgtgttcct	gccactgcac	17760
tccagcctgg	gcagcagagt	gagaccctgt	tggggggaaa	aaaaaaaaaag	tctttaactt	17820
aaataaaatt	gacatttaaa	atcttaaaat	atctcatctc	tgtttcagta	ctaactctgc	17880

atttattact	ttcttttttaa	taggactgaa	ggaattttct	ctgagtata	ttcattcttg	17940
ggaaaaagaa	aaagataaaa	atatttttct	gttgagtccc	aatttgctgc	agaagactac	18000
aatacaagta	atagcatggt	attgaatat	taataaaata	ctatttggtt	catatgattg	18060
ataataaagt	atgaagttcc	ttgtaacacc	ttgcattgtg	aagtgtatta	aaaacctgct	18120
aagagtaagg	aataaacttg	tttaaaatat	tttattctgt	aatctcttta	aattatctgt	18180
acaaattatt	gacttaacct	aaattttaaa	atgaatgcct	tagcacaatt	aagttccaag	18240
aatagagttg	atcatgttaa	ctggtaaattg	gatcatgatt	taaaattctt	ctaggattga	18300
aacaaatgaa	aacgtagttt	taagggtttg	atttttttaa	ttcctatttt	tacatgcaat	18360
tttactgcac	aacctatctt	attttgacag	ttcttaaatt	cgcaactctt	cagaaatatt	18420
atcagatcac	ttttctttgc	ttccataagt	ttttttatta	ttatattatt	attttttttt	18480
tttaaaagac	ggtgtctcac	tttgctgccc	aggctggagt	gcagtggcat	gatcatggct	18540
cactgcagcc	tcgacctccc	aggctcaggt	gattctccca	cctcagcctc	ccaagtagct	18600
gggaccacag	gcgaatgcca	tgatgcctgg	ctaatttttg	tatgttttgg	agagataggg	18660
tttcacccatg	ttgcccagaa	ttgtcttgaa	ctcctgggtt	caagcagttg	ttctgccttg	18720
cccacccaaa	gttgtgggat	tacaagtgtg	agccactgcg	cccagctatt	ctagaagtat	18780
tttaagagtc	atcttttttt	tttttttgag	atggagtctc	actctgtcac	ccaggctgga	18840
gtgcagtggc	acactctcgg	ctcactgcaa	cctccacctc	ctgggttcaa	gtgattctcc	18900
tgccctcagct	tccttagtag	ctaggattac	aggcgcatgc	caccatgccc	tgctattttt	18960
tgtagtttta	gtagagacga	gatttcacca	tgttggccag	gctgctcttg	aactcctgac	19020
ctcaagtgat	ctgcctcctc	cagcctccca	aagtgcggg	attctaagtg	taaaccacca	19080
caccagacca	agagtgggtc	ttttacaata	ttattttttg	attaggacat	tcattcttgt	19140
cataaaattg	aagatactct	agtcatttag	aatttcattg	ttttggaact	agacattgtt	19200
tctttatttt	tgaaatgtta	ttgaaggaa	accatttgga	gaagatacaa	atgtaagaat	19260
tgtgaaaagg	ataattgtga	cacaaatcaa	aattatagat	aaaaatatac	ctgtaaaatg	19320
tattaaggca	ataacattct	ttctgcttgt	tgaccataaa	tatttatatt	ccttggtggg	19380
gtacattgtt	attgtcaagg	gtgttttaaa	aatgatcttg	catgcataat	ttattctctc	19440
tggtataaca	gaatcagcaa	tttagttttc	tgggacccca	gaaaaacatg	caaaagacat	19500
actttgaaat	gtaaaactga	tttttccttg	caactgtagg	tccttctaga	tcctatggta	19560
aaagaagaaa	acagtggagg	aattgacttt	attcttcctt	ttttaaagct	gagctctttg	19620
gggtaaagaag	ttatggccaa	actagcatgt	tagacatggt	tttaacacta	tatctggcag	19680
agtttttcaat	gtaaaatatta	aagttagatgt	taatgtcaat	aagtgatctt	aataatgcat	19740
cagtagatat	tttttcaagg	attgtctcta	tcttcacgcc	tagcttataa	tttgccctgt	19800
cgtctttttt	tttttctctt	tatttttatg	tttttatcca	tccttggtgg	taggggataa	19860
ccttgtcttc	ttcgataaca	agaagtctga	agcttattag	aaattttact	ttgagaattg	19920
atcgatgaga	agaaagcaac	tagatatcac	gtggatcata	tatgcttgaa	taaaacaata	19980
attcttagaa	caaataaata	catttttaaa	gttaaaagcca	aaaacattag	ttgaatgttt	20040
aaaaatat	caaattaaat	tattccttca	ctgtcttgta	ttactgtaat	aatttggatt	20100
atttgtgttt	ttctcaactt	ttaaaacaaa	tattttaaaa	attcctcttt	tgattaagta	20160
gggctagata	aaatataaaa	aatatttttt	aaactcctct	taatttccat	atttcttata	20220
taatatgaga	atctcttata	aacactacct	cttagaagtc	tcacagagaag	ctttggtaga	20280
tgtagtagta	gggatttgat	ttcttagaat	ggtataatct	gtaaatgttt	tagtaaaagg	20340
attaaacgat	aaagtcaaaa	tgtttatagc	acagtgttta	ttaatataaa	ataaaatctc	20400
tttttttttt	tttgagatgg	actctcactt	tgctactcag	gctggagtgc	agtgttgcaa	20460
tctcagctca	ttgcaacctc	cgctcctggg	gttcaagcaa	tccttcgcga	tcagcctcct	20520
aagtagctgg	gattacaagc	atgcaccacc	acacctgcct	aattttttgt	attttttagta	20580
gagatgggg	ttcaccatgt	tggccaggct	ggtctcaagt	gatccgcctg	cctcagcctc	20640
ccaaagtgtc	gggattacag	gcgtgaacca	ctgtgcccag	cataaagtaa	aatctcttca	20700
gactctcatg	tgatcatgta	aagtggcagg	cagtcacagt	caagaagtag	tttaaagttc	20760
atgtttgttaa	aatataatct	acagattgat	actggatttc	ataggtaatg	tttaagagaa	20820
aataagtttt	tagttatcct	cagtacttca	aaagcaccca	tttatgatta	tggtgattac	20880
taaactaaat	catttggggg	ctagaggtgt	ttttttatgt	gttaagattc	cttaaggagt	20940
tctattaggg	caaaactttt	agtaactgca	tatttttaaa	gtaataaaa	taattttaaa	21000
agcttgaggg	ctgggcgcgg	tggctcacac	ctgtaattcc	agcactttgg	gaggccaagg	21060
cgggtgagtc	acttgaggtc	aggagtttga	gacgagcctg	agcaacatgg	tgaaaccttg	21120
tctctactaa	aaatacagaa	attagccagg	tgtggtgggtg	ggcacctgta	atcccagcta	21180
ctcgggaggg	taaggcagga	gaattgctcg	aacttgggag	gcagaggttg	cagtgaagccg	21240
agatcatgcc	actgcactcc	agcctgggtg	acagagcaag	actccgtctc	aaaaaaaaaa	21300
aaaaaaaaaa	gcttgaagtc	agattcgaca	ttaatcagta	tactttctct	caagtagggg	21360

acaattttcta	agatttttagt	cttttaaaat	ttattaacta	gtctgagcat	gggtggcttgt	21420
gtctataatc	ccagcacttt	gtggggccga	ggcagatgga	tcacttgagc	ccaggagtgt	21480
gagactagcc	tgggcaacat	ggcaaaaacc	cgtctctaca	acaaatgcac	acacaaaaaa	21540
cccaatcagc	tgggtgtggt	gttacactcc	tgaagtccca	gctactcggg	aggctgagcc	21600
aggaggatca	cctttgccag	ggcgtttgag	gctgcaggga	gctgggttca	caccactgcg	21660
ctccagcctg	gatgacacag	caagcccctt	tctcaaaaaa	aaaaagataa	aaaattaaat	21720
taaatttaatt	aactacactg	ggaaggcaaa	attcagcatt	tttttatagc	ttaaattttat	21780
cctgcttcag	tcttttatca	tgtaactatg	tatatTTTT	acagaggagt	gaattcctta	21840
ggcgtatcct	ccttgagca	catcactcac	agcctcctgg	gacgcccttt	gtctcggcag	21900
ctgatgtctc	ttgttgagg	acttaggaat	ggagctcttt	tactcacagg	aggaaaggta	21960
agtgggttaag	gtgtgttcat	ttttctgtaa	catttaataa	cttttcattt	atctttcttt	22020
gggttttgac	catctattat	atagggtggg	ttttgaccat	ctattatata	gggtttatata	22080
gacatatgga	aagcattcat	ttattcacta	atatttctgt	gtgtctgctt	ttagggtgtg	22140
ggggagtgat	gacgaataag	actgatgttc	tccatgccct	ttttctgtgt	cagttgatac	22200
aattatatgg	tttttctttt	ttaggctatt	agggtgtgat	agggttgagt	aacttacaaa	22260
tggtgaacca	gccttgcata	cctgtgataa	ataccacgta	gttgtgggtg	atcattcttt	22320
ctacattgct	gagttttatc	tgctaattgt	ctgttgagct	tttgtccatt	taagtttgaa	22380
agtgattagg	ttgcagtttt	ctgtttttgt	gtgtctttgt	tctgggtttg	ctatccgtgt	22440
aaatctggcc	tcataaaaatg	agatgggaag	tattctctcc	tctttttttg	tttttttggg	22500
agaggttgta	taaaattgag	gctgaatctt	gggtggttgc	acaatgacag	gaactatttc	22560
tgtgactgaa	tatatggga	attcctataa	agcaattatt	ttctagggaa	gtggaaaatc	22620
aacttttagcc	aaagcaatct	gtaaagaagc	atttgacaaa	ctggatgccc	atgtggagag	22680
agttgactgt	aaagctttac	gaggtatgag	tatggtaaca	ctctatataa	atcccttttt	22740
cattagaaag	acaggaatgt	tatacataat	gctgtcaatc	taataaatac	acatatcatc	22800
tagtctttaa	cttttctggt	tatcatttag	tcattaaaaat	ttctttggct	ttctaattgt	22860
tttgataaaa	tttctaaaac	tctccatatt	taatggaggc	ctattttttt	ttctagccag	22920
aactttttgt	agactacatt	tctggaagtg	ctcactgaca	ccactctgaa	aaattagtag	22980
ttagaatata	ctctaattgg	tataaatgat	ctctgaattg	ctatggaaaa	ctggggagaat	23040
ggttgcttca	ggggagagaa	agtaggaggc	tgtggacagc	aatgaggaga	attacagttc	23100
accataatac	acttttgtac	ttttaaagtc	cttaacattt	acattattat	ctattccaat	23160
aaaaataatt	gggaagattt	tactttgaac	agtaaatttt	tcccccatgg	gtaccgctgt	23220
catatagtct	caactaatca	tgaacttggt	tatttcctgt	tctttgtaaa	tttaaacttt	23280
gtaactcacc	aggaagtgtg	aagccaaatt	tgtgtttcaa	atatagcaac	tccaggatct	23340
ctaggcagat	gcattttgat	ttgattttta	atgaatcttg	atcccttact	ctcacttatg	23400
ttttcccaca	tcctactttt	tttattttgt	tgtaaagcat	ctaaaattct	caatgggatg	23460
aaactgggta	taaatgaata	catgcataca	ggaattatag	tagcatattc	cttttctttt	23520
ttcttttttt	ttttttttga	gacagagtct	tgctctgtag	cccaggctgg	agtgcagtgg	23580
tgcatctctg	gctcactata	gcctccacct	cccaggttca	agcaattctc	gtgcctcaac	23640
ctcccagagta	attgggacta	caggtgcatg	ccaccacacc	tggctaattt	ttgtattttt	23700
tagtagagat	gggttttcac	catgttggcc	aggctgatct	caaactcctg	acctcaaagt	23760
gatctgcctg	ccttggtttc	ccaaagtgtc	gggattacta	gcataagcca	ctgcacctgg	23820
cctccttttc	tgagttttat	aaaatttgat	actttactgc	acgctttgag	actgtattaa	23880
ttgaaccatg	ttgatgaaca	agtttttgtg	atgggtatat	taataaaaata	tagatcaaat	23940
ttttatagt	aaatcaatat	cgagcttttc	tagtgctttc	aaaaggacaa	cctgaatttt	24000
cccagcactg	aaatgatact	gaaaccattt	catatcttct	gtattaagga	aaaaggcttg	24060
aaaacataca	aaaaacccta	gaggtggctt	tctcagaggc	agtgtggatg	cagccatctg	24120
ttgtcctgct	ggatgacctt	gacctcattg	ctggactgcc	tgctgtcccg	gaacatgagc	24180
acagtccctga	tgcggtgcag	agccagcgcc	ttgctcatgg	taaatgcac	caccactggc	24240
tttaaggctct	gttcttttgt	cagtcagcat	ttttagtctt	aacaataaat	ctactctctt	24300
cagagaataa	tatatgtgtt	atgttaagtg	ttgtgtttga	ggcccctgat	ggcattctac	24360
agttgtccta	tagactgtaa	tagcaaaatt	ggtagagtaa	aaacagtgtg	aaaattctgc	24420
aacttcatgg	ttagtccctt	agggtttttc	attctccctt	acttattgtt	taattttacag	24480
ttttactctt	ttgttcattt	gacaaatatt	tgtcaaatgc	ttgtgcacag	tctgtattct	24540
caattcttag	gagaaaaaga	agggtgaaca	gtattagcgc	agaacgatac	taataatgat	24600
ggctactgtg	tatgagtagc	cagccctttc	ttggctttct	tggattgctt	tgtattctac	24660
atgaagatat	tccctgggct	ttacaggtca	ataaatggaa	attcagagag	attaatttga	24720
ccagggtgac	caacaaggag	atgacagcat	acactatgcg	agaagtatac	acagagtagt	24780
gtaggagcat	ataacctaaa	ctgggggtga	gggtgggataa	ggagttatca	gggaaggctt	24840

tttgaggag	ttgacaactg	agccgagttt	tgatggaaga	gtagaaatta	gcatgaacca	24900
atctcatgct	aataaagaag	caaaggaagc	gtggtctaca	ggcaaaagca	cagagggtaca	24960
ggaagtaatg	atatgttggg	gaataccctg	ttgactggag	cttagagtgc	aaggagagga	25020
gtgctagggg	ggtgaggttg	gaggggttgg	cagcattgac	ttgcttcaag	gttcttaaga	25080
gctgaaatag	atataaaatg	caactaagag	tggcttggat	tattattacc	tagtgtgtta	25140
atctcaaatt	ttgaaatcta	tagcatctat	aggactgggtg	ttactaatct	tacactcgat	25200
ctggttactgt	tcttatacta	gatctattag	tccagtgttt	aagggagtgg	tgcagatttc	25260
taggtcagga	caggactcag	atgtacatta	ttaatgccta	tttcagttct	gaccttctca	25320
tatgaaacct	tataagacct	ggggtaggaa	gagattgttc	tggaaagtc	aggaatatga	25380
actgtatttt	gtttaacaaa	caatacagta	tggaaattta	tcacccttcc	agaatattta	25440
tttcagagac	aaatttttat	cattcgttca	tttatttcat	aagatccacg	agtagggaac	25500
ctcactagac	attgctctga	gtatatggtc	tgagtttgca	gtacctcttg	tgtctccatt	25560
agatttatta	ggtcctcaat	agataaatca	gggaataact	agatggattc	attttttaaa	25620
gacatgaaag	agcgatacca	tacatactgc	accttaaaag	tcaaccttag	agtatcatta	25680
tttttaaatga	atgtataatt	tttaaatctc	atgtttactt	ttcctaagct	tttgactat	25740
attgcttaat	tccagctttg	aatgatatga	taaaagagtt	tatctccatg	ggaagtttgg	25800
ttgactgat	tgcacaaggt	cagtctcagc	aatctctaca	tcctttactt	gtttctgctc	25860
aaggagttca	catatttcag	tgcgtccaac	acattcagcc	tcctaatacag	gtaatacact	25920
acttgtaagg	attattgaat	tatgtccctt	ttatagaaat	tatttttcaa	ttttattagt	25980
aattcgtggc	tttaaattta	tgtttctctt	aatgatttta	aggatatgta	agtcaacatt	26040
tggtgcatat	tgtgctagag	gcataaatta	taatttatag	ccacctgaaa	tgtagtatg	26100
cgctttccaa	gaaaatgact	tttttgaaaa	tggtatttct	ttgaatgaga	aagaacagag	26160
agaaatagat	agatggcttt	taaacacttc	attaattaaa	cttttttttt	ccaccatcac	26220
ataatggcac	ttagtccctt	ttgggaactc	atgaggggtt	tagtggtagt	gagctgaaag	26280
aaatatgttc	caggactggc	aaacatatct	taaattcttt	aaaattttca	cctagcatct	26340
accctaaata	ttcagacctt	gtgctagtta	actgctattg	aagaacaaag	gtattatata	26400
tattattaag	gataatagaa	tggtatttga	gatattgggtc	attgaatatg	aatatgtttt	26460
gagaaataag	ttttatagga	acaaaaaaaa	aattcttaaa	ggaaccatat	attactaaaa	26520
atgcttctta	ttggagaaaag	aaatgacaa	catttatata	tgtgattttt	tcacaacttt	26580
attaagatat	aattttaagta	caacaaactc	acataaagtg	tacaatttga	tcagttttta	26640
catatgtaga	tgccatgaaa	ccatcaccac	aattaaggaa	acaaacattt	cctagcatct	26700
agaagtctcc	tagccctttt	actaccctat	cctccctctg	tccatcccca	gacaactacc	26760
aatttgcttt	ctgtcactat	agatttgtca	acctgatttt	ctccaaatat	acattcaaaa	26820
atatacagtt	gaatacaatt	ggaaattoga	attttgtgtt	ttttctttta	ggaacaaaga	26880
tgtgaaatcc	tgtgtaattg	aataaaaaat	aaattggact	gtgatataaa	caagttcacc	26940
gatcttgacc	tgcagcatgt	agctaaagaa	actggcggtg	ttgtggctag	agattttaca	27000
gtacttgggg	atcgagccat	acattctoga	ctctctcgtc	agagtataat	caccagagaa	27060
agtatgtttt	actattaaaa	cctgaacttg	gaatcttctt	tctattgtgg	agaaatgtaa	27120
ttgtagtaag	acaagaatta	aatatattcc	attgtagtat	ttgaataagc	agttatttga	27180
gtagaaaatt	agtgtttcca	gctaagatga	tggcatattt	tgaataattca	tatagtgaat	27240
ataactagta	aaagaagttt	tgtttatttt	taaacagaat	tagtttttaac	aacattggac	27300
ttccaaaagg	ctctccgggg	atttcttctt	gcgtctttgc	gaagtgtcaa	cctgcataaa	27360
cctagagacc	tgggttggga	caagatttgt	gggttacatg	aagttaggca	gatactcatg	27420
gatactatcc	agttacctgc	caaggtatgt	ttaaaaaaag	aaaaagtga	tacttactcc	27480
cagaagaacc	actgtattat	tggctttggc	tttatgtgtc	agcttgccca	atctccgtgt	27540
gagtcaacaa	gtgtttactg	agttaccaaa	taaatgtctt	aacactattt	taggtacttt	27600
aacaaatttt	aattttatta	attaattttt	tattagaatt	gagacctcac	tctgtcatct	27660
aggctggagt	acactcacag	ctcactgcaa	cctcaaaactc	ctggggtcaa	gcaatcctcc	27720
tgcctcagcc	tccccagtag	ctagaactac	aggcatgaac	caccatgccc	ggccaactct	27780
ttaatcttct	tagagacgga	gtcttgctat	gttgcccagg	cagacagatt	ttaatgtgta	27840
tgatgcagtc	tttgatgata	agaaacttat	aatggaaagc	tgaggtgata	gttacagtaa	27900
atacattttg	atgtataatt	ctgtttgctt	taatcattca	aattgtagta	aagcaagatg	27960
aactgtctgc	tgggatttga	gcagaaatgg	ataggaataa	actaggaggt	agaagagtta	28020
tcaaggttca	caggactgat	gggtgaagct	agattttccag	accgggagtg	tcagtccttg	28080
aaaagcagac	ttggcaggca	tagacgaggc	agatagcagg	ataaaggaga	caaagttaga	28140
ttgttcttca	gaagatcaga	tggtagagtc	taggaggtag	tgtgttttaa	tcagagatct	28200
gagaggcaaa	gatcattgca	tgagatcagg	gacccatgca	aaggagttag	aaaaaaaaact	28260
gggttaagga	gcctgctgca	tggcaactcc	tgggaacagt	ggccactggg	gcctggggaca	28320

tgttgattgc	agcccaggac	tgttaaaacc	agtgtgagag	aacatgggta	tggaagtact	28380
agctagcagg	atcatgaccc	cgatgctggg	atggggcatc	aagcattagt	acatggagat	29440
tcagtacatc	cagatgcagt	acatggagac	tatatgcgta	actgctgact	ttgggcttct	29500
ttcagattgg	agcagaggta	gaggtgagtg	ggaatattct	caatagaggg	aactaaatag	29560
gcatacctaa	taaaggagac	caggatattg	cagacagtag	cctcatgttt	ggctcacctg	29620
ttcaaaaagt	tctcttggtc	ttgagcagtg	gtgccttaaa	aggtaacttg	agaagcagtc	28680
gattatttgt	tcagcctgga	gactcttggg	atattttact	atctttgatt	gaatagattt	28740
aaatgtacac	agctctcata	acttgcccca	tgaagcatat	ccatgaaagg	cactatactt	28800
gttaaaaagat	tggtttgtac	tttttaaatg	tagtactttt	aataaaaacag	gaaaaataga	28860
agttctgatg	cagttatatg	cattttatat	agaatgtgtt	cttaattgga	aaaaatttgt	28920
cgtagttcct	ttgagttcat	ttacagtttt	tagtaggaat	tgtattttct	actgttgtac	28980
ttgctgttac	taaagaaaga	tggtcgtgat	taccatctga	attttttttc	tatacatctga	29040
tcttttagctg	ctacttagtc	atttctgttt	agacttgagc	tctttttcat	attttttttt	29100
tttgtttctc	agtatccaga	attattttgca	aacttgccca	tacgacaaaag	aacagggaata	29160
ctgttgatg	gtccgcctgg	aacaggaaaa	accttactag	ctggggtaat	tgacagagag	29220
agtagaatga	atttttataag	tgtcaaggta	tgttgctctac	ttatcttctt	tttttatctta	29280
ggtaaaaatta	acataaatgc	agttagccat	ttcaaagtgt	aaattcactg	gcattttagt	29340
cattcacatt	gctatgcaac	caccacctct	ctctaatttc	aaaacttttt	cattccactg	29400
ctcctcttgc	ttatcccttg	gcaaccattc	atctgctttt	tgtctctatg	gattttgcctt	29460
ttctgtatat	ttcatataaa	acaaatcatg	caatatgtga	ccttttttgt	ctggcttctt	29520
tcacttatgt	aatgttttca	tggttcatcc	aggtagtagc	atgtatcagt	acttccattcc	29580
tttgcatgac	tgaataatgt	taccatactt	tgtttatcca	cttatcagtg	gtgaacattt	29640
gaattgtttc	taccttttga	ctattatgaa	taatgttgct	gtaaatattc	atgcacaaat	29700
ttctccacgg	atatgttttc	atttctcttg	ggtataaaact	gaggagttaga	attcctgggt	29760
cttagggtaa	ttctctaact	tttcaaagaa	ccaccaaaact	gtctttcaca	ccaactgcac	29820
cattcccaact	agcagtgtgg	ggggttcctg	attctccaca	tctttaccaa	caccattatg	29880
tttctcaatt	gtgggctagt	ctcacatttg	gaaagctagt	gggagcagcg	atccatctat	29940
taaaaagttgt	atgaaattga	gtaatgagcc	acctctctct	tgtagggtct	attatgttct	30000
tgcttaaggc	aatcttcatg	cattgtgaac	agaattatac	ataaatgctc	agataaaaagg	30060
gcaaacattt	cttaaaaggga	gtagacaact	agaggcagga	gaccatactg	aggcaggaag	30120
ctggggtttt	tatggttctg	ttacttttga	ctatatctca	ccattgcttt	tgtcaaagtg	30180
agactaggtc	taagtttttt	tcaggtataa	ggtgagtgtg	gtaattaaagg	ggcatgctag	30240
cagatcattt	tggttaaatgc	ttcacagtcc	accactgggtg	tgtcattgtg	gtcgcagatc	30300
cagtatctta	gctgtgtaat	ttcagacatc	agcaatatta	gtttaacaaa	gggcaattag	30360
attccaagac	aaaggaatcg	tgtattattc	tagccttatt	caaacttgat	ttataaatca	30420
gtttgttaat	ttattttatt	gtttctgtat	ttatttttat	ttctttgaga	ttgagttcca	30480
ctctattggc	caggctggag	tgtagtgtatg	caatcttggc	ttactgcaac	ctctgcttcc	30540
tggtttcaag	ctattctcct	gcctcagcct	cccgagttagc	tggtgattaca	ggctaatttt	30600
tgtattttta	gtagagatgg	ggtttcacca	tgttgccag	gctgggtctg	aactcctgac	30660
ctcgagtgat	ctgcccgcct	tggcctccca	aagtctctggg	attacagacg	tgagctaccg	30720
tgcccagctc	agtttagtaa	tgtataactg	ggttttaccc	agttgtaaat	tactcttttg	30780
tcgtgttttt	ttgagaactg	gcaatgacgg	agaaaactaaa	agtgccaggc	tgttgccctg	30840
ttcctgttat	tttgcccttag	tttttttttt	tttttttttt	ttctctgaga	ctgagtcttg	30900
ttgtgttacc	aggctagagt	ggagtggcat	gatctcggct	cactgcaacc	tctgcctcct	30960
gggttcaagt	gattcctgcc	tcagcctccc	gagtagctgg	gattacaggc	gcctgccacc	31020
gcacccgggtg	aattttttgta	tttttagtag	agacgggatt	ttaccatgtt	ggccaggctg	31080
gctcgacct	cctgacctca	tgatccacca	gcttcggcct	cccaaagtgc	tggtgattaca	31140
ggcgagaacc	accgtgcccg	gtcttgccct	agttatttct	tgttccctcc	tctagtccta	31200
tagttctctg	actgtattga	ggaaatgtaa	ttaaatatta	ttatgttaat	agatatttat	31260
gtgggtgaat	attagaaatt	ccttattttg	gtcacatata	ctgatcagta	gttgggtctt	31320
tgtagatagt	gattttttcac	tagagatgac	tttaggacct	attcagggtt	tttttaagat	31380
cccaatttaa	ggaaagacta	ttctcattat	tgattttgct	atatgcaggg	aaatttatct	31440
cgaagggttt	ttcagttggc	ttttagggaa	gattatatat	tctctttttt	tttttttggc	31500
cttttccac	atgttctaaa	aatgatatat	tctttaaact	ctatgaaaat	acattgtttc	31560
agtaattgaa	gatgctgatt	aaagtcatat	ctctacacat	tttttaaaat	ttgagataga	31620
tgggactttg	tcccttctta	caccattcac	ttattcactt	ggaaaaacta	ttatccaata	31680
cttatgtggc	agacactggt	tctggcacaa	gggattcagc	agtgaacaaa	actgcctttt	31740
tggagtttac	attctactag	tggaaagcga	caacaagcag	atagacacat	tcagtatata	31800

attcactgtc	agatgggtggt	ggtaagtcct	atgtaggaag	aaaagcaggg	taaggaggct	31860
tggagtaact	ggagtgaagtc	atagatggac	ttgtcaggaa	agggtttctg	aagagggtgg	31920
atttgggcag	agatctaaat	aaaatgaagc	aacaagccat	gagaatatcc	gggggaaaat	31980
gttctgggca	gaagcatcaa	gcatagaact	tgtggtatga	tattttattct	agcacacatt	32040
aattttaaaa	atgtataaaa	gacatccatt	taatcatatt	aaagatttcc	atgattcatt	32100
tagacttaag	cagaaaccaa	atttatattt	tctttttaaa	taattttatc	tcaactctta	32160
ttttacccaa	tagggggccag	agttactcag	caaatacatt	ggagcaagtg	aacaagctgt	32220
tggggtatatt	tttattaggt	tggtagccta	tgaatgtttt	taaagtaact	gactctgtta	32280
ttattttatca	atcagtgcct	tttttggtct	tgttttttga	agaactgata	tttgaaacct	32340
gtgggtttatg	tgaattatta	ataagctaga	ggacgtggat	tctctatttc	atcaataaat	32400
acaaaacatt	ttagatatta	aatttttgaa	attattttgt	tttgttttac	aatagaaata	32460
ctcctcaaag	tggaatcgaa	gtggttattc	aaagaaatct	cagagtagat	tcttatatga	32520
agcaataaat	tgcccctaatt	ttatctctaa	attttgtaag	ttctaaattc	ttttttcccc	32580
cagtttctaa	tttatctctt	ataagtcaag	agtccatctg	gccaatttaa	tttcagttag	32640
tgtaactatt	ttgcatatat	taaaaaactg	tatatgaata	cagaagatgg	tatttaagga	32700
tgaaaataat	tattcaaattg	tgatagcatt	atggggaggt	ttaaaaataa	agttactgtt	32760
ttattcttcc	aaaaatttta	ttataaaagta	tacagttaag	agaatataca	taaaatacat	32820
atgcagctta	aggaagaata	ataaaatgaa	tacttcatgt	attcaccacc	gagtttacca	32880
ggaaaaagca	taaacaaaaat	aaacctcttc	cacgtaattc	ctgggttaaa	gagaagttat	32940
agtggaaaaat	atttggggagc	aaacgataat	gaaaaatacta	tccattaaaa	ttggttagatg	33000
ttgcaaaaact	gatttcaagg	aaaatttata	gtgttaaagt	tttagaaaag	aaaaaagggt	33060
agaagttaac	cacttatgta	tctatctcat	gaaattagga	aaattataga	tataaactaa	33120
aaaatatggt	aaaagggaaa	taataaaagt	aagaatgaag	tttaatgaaa	cacaaaaacag	33180
agaagctcac	aaagccaaga	tttatttttt	gaacaccgag	tacaattgac	aaatctctaa	33240
caagtttgat	taagaaaaaa	gaaagcatga	ataaacaatt	ttagggataa	aaagggaaac	33300
atcgctaaag	atatcccaga	aatgtaaaag	ataataaggg	aatattatga	aaatattcat	33360
gccaatacat	ttgaaaactt	aggtgacata	gacaaaaaca	aaattgacca	aaattgagca	33420
aaaaagaaac	aaaatctgag	tagtcctgta	acttagtaaa	aattgagtta	gaaaagttaa	33480
agaagtcctt	acacaaatca	aacatcagac	tcagttttct	aggagagttt	tgccaaaacat	33540
tcaaagtagca	tataattctg	gtctattttt	ggccccagaa	gatatatttt	acttgccatg	33600
catttaatatga	gatagctggt	gatttttttc	aatcacctgt	acaggtgttt	tatattaggt	33660
gttatttcgcc	agacatctag	tccacctgtt	gccagatatg	gaattaatat	tcacttattt	33720
tgaattaaaa	tttggttaata	aattaataaa	acaaagtcaa	agttcaaatt	attaaaaaag	33780
taaaagaaat	aaaatatatt	ttatagagag	cccttacaaa	acagtaccaa	cataatgagc	33840
tttccaaatt	ttgaatgggc	aaaataaatg	aataggcatt	tcacaaaaga	aggaagggtg	33900
gccaataaag	atatattaat	ataaaaatgg	ttacttgtaa	taggaatcaa	aagtgtttga	33960
cttattgact	aagagtcagt	ttttgttttg	atccctgtta	gtctatccag	aaggcatggg	34020
tcttaataaaa	caccttgacc	tcaacagttt	actgaataca	agggtaattt	catatgcctt	34080
gccttcttta	agggtttggt	gtaaagatta	aaataaatac	ataaataatat	ataaatacat	34140
ttatatgtat	ttatatgtaa	ttacatacaa	cttgccctct	ttaagggttt	gttgtaaaaa	34200
ttaaaagaag	tataataata	tataataata	cataaaataa	atacattcat	atatgtatat	34260
gaaatcactt	tgccaactat	gaagcctgat	tcaaataatga	aatgttgttt	gtttttccca	34320
gagcacaggc	tgcaaagccc	tgcattcttt	tctttgatga	atttgaatcc	attgctcctc	34380
ggcgggggtca	tgataataca	ggagttacag	accgagtagt	taaccagttg	ctgactcagt	34440
tggatggagt	agaaggctta	cagggttaata	attataaata	cagaaataga	atgttataac	34500
aaaatgtcat	catgtcatca	gatttttggt	aaaaaatggt	cttttttccct	ctagggtgtt	34560
atgtattggc	tgctactagt	cgccctgact	tgattgaccc	tgccctgctt	aggcctggct	34620
gactagataa	atgtgtatac	tgctctcctc	ctgatcaggt	gacaatttca	tatttagagt	34680
ccaaaaccca	acaaatgcta	cactctttcc	ttgtgagctt	tacttctgcc	aggtaatggc	34740
aattgtcctt	agaagaccag	ctttcttagg	gaaaagcttt	agccactggt	tgctcaaagc	34800
ataaaaagat	tctgaattag	atgcaaagcc	tttttttggt	ccagtgcagg	tctgaaaact	34860
ttgtaatcct	tctgtgttgg	ctgattgggg	aaaaaaaaat	gcaagaaacc	taatgtatta	34920
tattttcaca	ttatcttctg	ttcaaagatt	acatacttcc	attatcctgt	caaaaaaaaa	34980
actctgatac	agaatcaagc	atgtgaatcg	taagcatgta	agcaggtttc	atagagataa	35040
tttttcaact	cttcttctgt	ctgtgttgtt	ccaactctta	ttctccaatt	tagaagcaaa	35100
caaataaatg	aatgaaagaa	cagatagaca	aatgaatagt	caaaggtata	aagtatctgt	35160
atatatgtta	catgtagcta	ttatttaaat	tatttagatt	ttccttttga	aataccttct	35220
tggcacactt	gcctaaatct	agaaaaataag	cactgtgtga	ataagaaatt	atttacactg	35280

aatattttgt	aggtttttgg	gtttttgttt	ttcagacaag	gtctcacttt	gtcaccacag	35340
ctggagtaca	ctgggtacgat	cacaactcac	tgcagcctct	atggcccagg	ctcaagcaat	35400
ctccccacct	cagcctcccg	agtagctggg	accacaggca	cacgctacca	tgcccagata	35460
attttattat	taatttttgt	atagagatgg	ggtctccctg	tgttgcccag	gctttcttga	35520
actccagggc	tcaagtgate	ctcccacctc	aacctcccaa	agtgttggga	ttacaggcgt	35580
gagccacat	gcccagcctt	aagagtgttt	gattttcatt	cattttccta	tatatattat	35640
ttctgttggg	gaaaaaatcc	caaggaagat	aaatagtagg	ctgttggtac	atttctcaac	35700
ttacttataa	agcttttttag	atatataagg	ttaatttatg	aagaaaatca	taagatacac	35760
aatttaagat	aatatttttta	attttatttt	ttatttgtta	aataaatttt	tctcctttca	35820
ggtgtcacgt	cttgaaattt	taaatgtcct	cagtgcactct	ctacctctgg	cagatgatgt	35880
tgaccttcag	catgtagcat	cagtaactga	ctcctttact	ggagctgac	tgaaagcttc	35940
actttacaat	gcccatttgg	aggccttaca	tgggaatgctg	ctctcgagtg	gactccaggc	36000
aagttatatg	aggaagtgtg	tatgacattt	tatgagtgat	aaaagaagta	caatgtcaaa	36060
atttccacct	taaaaaatgc	tatttttttaa	acaactttgg	taaaactgta	tagaaacata	36120
aattttacctt	tagttgaatg	ttcccatagtt	ggaatatggg	ttttgcagag	aattttataat	36180
tatgaagttt	gatgtctgtt	tctttaacat	taccttaata	ttggcaaaaa	catgttgggtg	36240
tttgcaagga	tattatttaa	attgggatac	catgaattaa	atactacaaa	caaaaaataat	36300
tagagttttt	tgtttgtttg	tactttaact	tttaaaaaat	aatcagttaa	agttgttgtt	36360
ttgaagctca	cattgttcca	atctggccaa	taggagcccc	ttttgtatgg	ctcctgtatc	36420
tttatgacat	gtcctcatca	ttcttgaatc	acttccctac	ttccagatac	agtaagttat	36480
tcttggccag	gtgcagtggt	tcacgcctgt	aatcccagca	ctttggcagg	ccaaggcagg	36540
aggatcattt	gggcctagtt	tgagaccaa	tcagtgttgc	acaaactgta	cccactatgg	36600
acaacagagt	gggatcttgt	ctctgtgaaa	aattttaaaa	ttagctgggc	atggtggcac	36660
atacctgtag	tccatgcttc	ttgggagagg	ctgtggcagg	aggatcgctt	gagtaaatcc	36720
aggatgcagt	gagccatgct	tgtgccactg	cactccagca	tggatgacag	aatgagaccc	36780
tgcccccaaa	aaagaaaaat	attcttggtt	tatcttgtac	tttctgtatc	ccagccctag	36840
catcagcctt	ttctctaaag	acagtattat	gattttaata	tttacagtag	atatttgaac	36900
tgttacatta	tagactttac	catatatttt	ctaggaagga	ttattctatt	actcttcttt	36960
accacatttg	tttggaatgt	ctacagaacc	tacagtttct	aaatcagaaa	ctccctagg	37020
ttttgctatt	ttggcaagcc	attgaagtcc	ttccctctcc	ctttactacc	agaaagggtg	37080
gtattttag	agctctctat	aatgagaaa	cactctataa	catggttgat	tcattctttt	37140
ggagtagaaa	agtatgaatg	gaaagtcaga	gacataaaaa	taaagcccag	aggtctgagt	37200
cttagcttca	ttacagactt	tcttggggga	tggttggtaa	attatctaca	cattctatct	37260
tgtctttata	attttaatat	ttaaattttt	accatgtgcc	tcaaaaccgt	tagagaatta	37320
atgagctctt	tgaaaaatgc	ttctaagttt	cttgtattgc	tctaatagaa	tgctatctat	37380
gttattattt	atttctgaga	ctaaaattgt	ttacatcttt	aaactggttg	tccttttgtg	37440
tatttttagga	tggaagtcc	agctctgata	gtgacctaa	tctgtcttca	atggctcttc	37500
ttaaccatag	cagtggctct	gacgattcag	ctggagatgg	agaatgtggc	ttagatcagt	37560
cccttggttt	tttagagatg	tccgagatcc	ttccagatga	atcaaaattc	aatatgtacc	37620
ggctctactt	tggaaactct	tatgaatcag	aacttgga	tggaacctct	tctgatttgg	37680
tatcttgtgc	agtcattcatt	atacagttct	gaaatataaa	gctatatgtt	ggtgtaaagt	37740
tgcagtgatt	tctctcctaa	ccagccccac	atattcttcc	tggttggttg	gttcttcagt	37800
aaaatagtct	tgtttcttgc	ttacactaat	tggttaattg	cattccttgt	taagattttc	37860
aagacagggc	tgggagcaag	gaaccaaagt	agcgctggtt	tgtgattacc	tttggtttct	37920
ttgaggtttc	tcttacctag	tggctttaaa	acatctttag	gagcagttcc	attttatagt	37980
aaacttaaat	tctgttatca	tgaacagttg	aggataatga	ataatttgat	acaataatgt	38040
aagaaattcc	tgaaaacaaa	gtgttatctg	tgatactttt	gctgcatagt	aagcacaatg	38100
aagtgtactg	ataatgtttc	aacaggaaa	tgttttgatt	aaatgtgggc	agtatcactg	38160
ttctactagc	attcaacatc	tcttctaaaa	attaatagt	gttcaactga	attttatggg	38220
tacatgtaac	atctgtacat	gtgtttggtt	atctatatgt	ttcctggttt	tttgtagatt	38280
tgctttatta	atttaggctt	tttttttttt	ttttttttga	gacagtctca	ctctatcatc	38340
cagactagag	tgcagtggca	caattatggc	tcactgcagc	cttgacctcc	tggttgattg	38400
tgattcttcc	acctcagcct	cctgagtggc	tgggactaca	ggcacatgcc	acctgcccc	38460
gctaattttt	gtatgttttg	tagagacgag	gtttcaccat	attgcccagg	ctggtctcaa	38520
actcctgggc	tcaagctatc	tgcgtgcctt	gacctcccaa	agtgttagga	ttacagggtg	38580
gagccactat	gcctagccta	actcagactt	taaaaaatata	aaagcaattc	atttttattc	38640
ccaagaacag	taaggtgggtg	gtttaatttt	agtctttaat	tctgttttta	atttattcta	38700
tttagaaatg	tcccagaaac	ttagtataac	tttactttct	gaaaatgaag	aaacctgtcc	38760

ttgggcatta	gtgtgttgga	tttaagcaac	aaagttaaaa	aaacctaccc	tgtgttatgg	38820
caattttcac	ttgatgggtg	ttctataaca	caggtatcag	tgaaccttta	taaaagatga	38880
acaacttttc	agcttgctta	atttcagtta	attaacatgt	atacttatct	atgttaatgt	38940
tttattgctt	aaaatgttta	atttttatat	ttggtaaaca	gatagttttt	tctctcccc	39000
tcttccttcc	atctttcatt	actacaattt	accatgcaga	gctcacaatg	tctctctgca	39060
ccaagctcca	tgactcagga	tttgccctgga	gttcctggga	aagaccagtt	gttttcacag	39120
cctccagtg	taaggacagc	ttcacaagag	ggttgccaag	aacttacaca	agaacaaaga	39180
gatcaactga	gggcagatat	cagtattatc	aaaggcagat	accggagcca	aagtggagta	39240
tggctttttc	cccctcatta	taattgttaa	aacttcttaa	aaattgtttc	acccttttga	39300
tatatatttc	tttgacttat	aaacgagcta	tatttataaa	caagggacca	gaacacatta	39360
actcagtc	ggttatgtgc	ttccttgctt	tcaatgtttc	attatcttat	aaggaagaga	39420
acgtatggc	tcttgaaaaa	actgacaata	agaagtaaca	actggactac	cacatttttt	39480
tttacatcct	taatttaact	cttcgtcaat	tctctttttt	acttaaggag	gacgaatcca	39540
tgaaccaacc	aggaccaatc	aaaaccagac	tggcttttag	tcagtacat	ttaagtactg	39600
cacttggtca	cacaagacca	tccattagt	aagatgactg	gaagaatttt	gctgagctgt	39660
aagtaacaga	ttctgttttg	gaagtacagc	tactattaca	agtgcacatg	tattacactt	39720
aaacctttta	agttcgtggt	taaaataaaa	atattttgaa	tatttaaaag	ctaattcaaa	39780
aaatatgtgt	cgtagctatg	cattaaaaaa	ccccaaaatg	tcagaagtac	agaagtcaaa	39840
attgagtttt	cattaaccag	ttcatttgat	tatatttgaa	ttattcataa	tggactcatt	39900
taatttttagt	aactttgggc	tgggtgctgt	ggctcatgcc	tgtaatccca	gctctttggg	39960
aggccaaggc	aggtggatca	cctgaggtca	ggagttcgag	gcaagcctaa	ccaacacggg	40020
gaaaccccat	ctctactaaa	aatacaaaaa	ttagccagg	gtgggtggcat	gtgctgttag	40080
tcccagctac	ttgggaggct	gagacaggag	aatttgcttg	accagggagg	tggagggtgc	40140
agtgcgcga	gattgcacca	ctgcactcca	tccagcctgg	gccacagagc	gagactgtgt	40200
ctcaaaaaaa	aaaaaaaaaa	atttagtaac	ttcgaagaaa	taagaaggaa	aattaaaagt	40260
tgaagtgat	tctaattgat	agtttataaa	attttgttat	aaaaatacct	gttttgctt	40320
caaaataatt	tatattaata	ttttattgac	ctcaagaaca	tttaaataca	ttcagattta	40380
ttcatttggtg	gaccacattt	gttatacatt	ggatttaaa	gatccttgca	attgagttta	40440
tggccacct	tgcatctgag	acccatggac	tgggaacct	tctaggtcaa	tgattcagtg	40500
tgattcaatt	taagagatgt	ttattcctgg	tctttagaag	ctgctacctt	ttgttatcta	40560
attttgag	actttgaagt	atgtatgtat	gtgtacatac	gttagtgcta	tgtatttatt	40620
aaagaagaat	cagaaaacag	aggtaaggaa	aaataaggaa	acaaatttct	gttaagccca	40680
ccacctccca	aagcatattt	gtttatatgc	ttatatatgt	tttctatta	tggtagaac	40740
agtctgtaca	tattgctata	tagcagtc	cctttatcca	catacatcct	gaaaattgtt	40800
ttacatttta	aatgttaact	actttattgt	ttttaaatgt	cattttatag	tgtagctatg	40860
ccacaatatc	caatttttag	acattttaat	tgctcccagg	caatgtggta	atgaacattc	40920
ttgcagctga	atatatgcac	atatctaatt	gtttcactag	gatagagggtg	gaattgtata	40980
acagggagct	cacatttttt	aaggcttttg	aaatgtattg	ccaaattgcc	tgccagatat	41040
actgcaccat	cactaacatt	gtgtgttgca	gtatttttct	aaacttggcc	cttttgattt	41100
tagaaaaatg	atatcaataa	tttacatttc	tttgattaaa	gtgtagaagt	tataattttt	41160
catattatc	attgtcattt	gtattttatc	ttttctaact	tgtctcttca	tcccctttgc	41220
tccgttttct	attggagtgc	aacttttatt	gtaagaattc	tttttaattt	ctgtgactgg	41280
aatttttttt	tctagtttgt	tatttcccgt	tcatttctta	aaatataatt	gtgtttgcca	41340
acaatccatt	atcttttgtt	ttgtaatggt	agtattttata	catattaaat	tatctctttc	41400
ttttttcaga	tatgaaagct	ttcaaaatcc	aaagaggaga	aaaaatcaaa	gtggaacaat	41460
gtttcgacct	ggacagaaag	taacttttag	ataaaatata	cttctttttg	atttggttct	41520
gttaagtttt	ttgatggctt	ttccatattg	tgtaacagga	aaaaaatgg	gtctatgaat	41580
ttcttcttaa	tttaacaaat	ttggtttaatt	tataaaatca	cagattggta	aatgtctata	41640
ttatgtaatg	atcaggattg	agattaatac	tgtagtataa	attgggacat	tataacagat	41700
tccatatttt	atttcctaaa	atctaaattc	agtctttaat	gaaataatat	tagccaaatg	41760
gtggaactaa	tttatttctt	ttgaggaaaa	gataataaag	aatgtaatta	aattttaaatt	41820
tcttggaatt	cccagttgta	tattcatcac	ctttgtagca	tttgacaaat	tttatgctta	41880
gcagcttctt	cactgttttg	aaataaaaata	tctattacc	tactgatata	attatctgtt	41940
ctttgtatat	caaaaaatgt	gaaatttaca	cataattcaa	atacatttaa	ttatccgctc	42000
aaccagaaat	gaaatcacat	ccctctacta	tactacatcc	agctccaagc	ccaagatatt	42060
taaatgacat	ccattcctct	cctagtcca	gttatgattt	tatcttgata	ttctctcata	42120
tatgaactaa	attataaagt	tagccaccat	caatacaatc	tgcgtatcta	atatcttaac	42180
tatatagtaa	tggggtaagg	gaacagcaaa	aaggagaaca	ttaattaaaa	tatacaagta	42240

agcctgggca	acatagtgag	accccatctc	ttaaaaaaaaa	aattagccat	gcatgatggt	42300
atgcctctag	tcccagctac	ttgggaggct	gaggtaggag	gatcacttgc	tcccaggagg	42360
ttcaagggtc	taaaccagca	aagctcagaa	tcccaggagg	tagaaacaaa	gacttagtgg	42420
atcactagta	ttaaactgag	acacgtcacc	ctgcattgca	ctttgtttct	cagttctttg	42480
atgaaatcac	tgagctgaca	tacctgccct	cttttcacca	taaagtgagt	ttcatgatca	42540
gaagcaatgt	ctatgggata	gcctaacaaa	caatgtaaaa	accatttagt	aagttcatga	42600
aggggtgggtg	tggtaaaaat	ttggagaaca	tacaaaacaa	atacaattcc	aagggtgtgc	42660
ccctccagga	aggacaaatt	gctgcctgct	ctgtgataga	agaggatcag	atgtaatcaa	42720
cctgccgtca	gacttgggct	gttctctcct	gggtgtggac	ttgcctgggt	ggtcactgct	42780
gctgacaagt	aggctgtcaa	tatagctggg	ttgtcatgtc	agctgtgggt	agggggaagt	42840
ccacattgtg	gaggccacat	ccctgcactc	ttggccaatt	tgaccatgaa	tcttaagcac	42900
tgggtgggct	ggaaaagaca	gccgattgac	atccatcacg	aggtcatctt	gaccacttga	42960
ttagtataag	cactgaaggc	ttttaactga	gcattcacat	aggacacaaa	tattctgatt	43020
ctttgggccc	attccaagaa	ctctgggcat	acttttcctc	cagacctcat	acccagttgt	43080
gttctttcca	aatttctggt	catctgggta	tggtattagc	cactatctgt	gaatcagcat	43140
agatttttat	atcagacatc	tctacctcct	gacagaatgg	aggagatatg	ttacttaaca	43200
attctgttcc	cttgggaagt	ttcctgtctc	cactgtttgt	aagggtact	ccctcaatgt	43260
agcagtaatg	ctttcactct	gatgggaagt	cacagtggaa	ttctgggtct	ccaagaatta	43320
gtgttagtgc	atacacagtg	tctgataatc	cccagagtgt	ctggtgccct	tggatcctgt	43380
gaagaaggct	tggagaaaag	aagattcatg	gcaagaactt	gtgatgtgat	gacagggcct	43440
tttctctggc	tcttcattct	tagtctgacc	taggtgtgag	aattaggtca	ggggccatga	43500
ctatatattg	gtgactcaaa	ccaggccttt	gtttactaac	tgggagattt	ttacattgta	43560
agaatcaagt	aggatctttg	cccattgtatt	ttggtcttaa	gaacacaaat	gatattggctc	43620
caatgactgg	aggaacacca	gggtccttgg	tctcacgctg	atttagataa	aacgactgtc	43680
aggcctctga	gccaagcta	agccatcctc	ccctgtgacc	tgacagtata	catccagatg	43740
gcctgaagta	accaaagaat	cacaaaagca	gtgaaaatgg	cctgttcctg	ccttaactga	43800
tgacattcca	ccattgtgat	ttgttcctgc	cccatcttaa	ctgagcgatt	aaccttgtga	43860
aattccttct	cctggctcaa	aacctcccc	actgagcacc	ttgtgacccc	cgccccctgcc	43920
cctaagagaa	aacccccttt	gattataatt	ttccactacc	cacccaaatc	ctataaaatg	43980
gccccacccc	tatctccctt	cgtgcactcc	tttttcggac	tcagcccggc	tcagcccagg	44040
tgaaataaac	agccttgttg	ctcacacaaa	gcctgtttgg	tggactctct	tcacacggac	44100

<210> 64

<211> 16869

<212> ADN

<213> Homo sapiens

<400> 64

aagcttttagt	agagatctca	aaaatgggtg	gatggtagca	aattactaag	aactctcaaa	60
gtttctaaaag	ccttagtttc	agcttgctag	aaaacctatg	ttgagtatta	tggctagttc	120
catagttgag	ttgggaaatg	tctttgagga	gacacttttt	cactttgtat	tcatctgtac	180
attttctggt	acttgcattc	tgatcatgtc	aggctattag	agcaggtaca	tttttataac	240
tggaaatggtt	atgtgtagtg	aagctctgag	aggactttgc	attagatctc	agcagcataa	300
tcagaagggtt	gtcctttgtc	tcagcaattt	ttaagctaatt	agtagcagaa	attgcagtgg	360
aaatagactg	ctttgccaca	acattcagaa	aatcatttat	ctttttattg	cagttcttgt	420
caccaaacaa	tacatttttag	tacttctcaa	attgcagaac	tctcataggg	ctgggaaaaat	480
gcctgtagac	acatacatat	tatgaatgtg	ctaattgttt	ttgtattttc	atagcccatc	540
aaagctcctg	agtcagtttc	cactataatc	actgcagaat	caatcttcta	caaggttaagc	600
ttttgtagag	ttactgaagg	aagagttggg	cctagtgggt	aatgtgccac	taaaatgttg	660
gattagtcta	aaggctctctg	ctactcttta	tttgataaag	gtgtgattat	actttttgtt	720
cccttcttag	ctgttttccc	ccataagtgg	ctgttattaa	aacatctcat	ctagagctga	780
agtgggagga	gaaagtgcct	actgacacat	gatgtgagga	tcttaagtat	tttttttttag	840
tgtagattgt	aggaattatt	cttaaaatgc	tgattgtata	gtgtggagcc	atggaagact	900
gagccgttag	tgcgatggca	ttgaagaatg	agaaggacag	agacaggatt	tggactagta	960
gaggttgtcg	actgtgtgtg	caaatgggta	gagtagggcc	agagattcta	aaatgccttt	1020
aagtggagtt	gagctgagta	agggcagtag	tgaggattaa	cacctactag	aaattcatag	1080
tgagaggaat	tccaagatgt	tttgataaaa	gaatgaggag	gtcagggttc	ccagggccaa	1140

agtccatgaa	catctgatac	ctcagtgaga	gaagtgacag	attgtttgtgt	ttaaaccaga	1200
agtcttagga	aaggaattag	aacatagacc	cccaaggctc	ggcaggcctg	gcacggcaca	1260
ggcagcaacc	attgaaggct	atttggtgtt	tcgggatctg	aactgtcatt	taggggacag	1320
tgggtgtgagt	tagtacttta	tacttgaccc	aggtggactg	agaaactcaa	gtgatgatgc	1380
ccttaagtat	actttttttt	aagcccacaa	tctatatagt	cgaagtctgt	tcctcccaac	1440
aggggtacac	tggcattcct	cagcagggct	gggaaaaacc	aacaacaaaa	aaagtctgta	1500
cacaggcaaa	catctctctt	atttttccaa	catttaatac	attgttaata	aaatatctaa	1560
agtttagcaa	acagttgctg	tgtatcagtg	gctgagcatt	ttgcatgctt	tatttcattc	1620
agttcactct	atgaggtgga	tactactatc	cccattttct	agatgagaac	attgaggcac	1680
agcgagggtta	attaacttgt	ccaagatcac	atagccaaca	agtcattggg	tgaggcagtc	1740
tcatgccaga	gcttaagcct	agagcatagt	tcctggctct	acagcttttag	caagtgactg	1800
gctatgtgac	gaggaccaac	ctctctaagt	tctcatctgt	aaaataggaa	ttgtaaatag	1860
ttactacctc	agtgggtcaa	atgaaatcat	atgtgttaag	cacttagcag	agtaagcact	1920
caatgaatag	taggagttat	cacatcttcg	tatttgtgca	ttaccttcac	agtttacaga	1980
ttaaggccag	aagcaacttg	ttgagctacg	ggttttagtgt	actaacagtt	tcctatgtgtg	2040
tctccatgga	aggggtgtgtg	ggacctgtta	ttgtgactgt	ctgtactttc	gtattgttgt	2100
ctgccaccca	tgtttattaa	atgataagga	caataatgca	acaaagtagt	caagtaatgt	2160
tgcaaatgcc	cagtattgta	gtggctatca	cagcagtgcc	actggcaggc	agcaccattg	2220
tggcaagttc	aagaggtcac	tgccagccac	tgagctagag	cccagatcag	gcattgcaaga	2280
ggagcctgag	tgggagccac	tggggatcac	ggccaagagt	gtgaccaccc	aagaccacga	2340
atggctgagt	ggcctccctg	gagcatggca	gtggcagaac	aactccatga	actcagatct	2400
gggtgatgcct	aaactagtgc	tgttctcgtg	tggaacctt	ttctctacca	gaaaccttga	2460
atcctctcag	caaatgagga	gactactcag	atcagtgact	tagtcctgtt	tggtgttata	2520
tatgtgtaca	caacacagca	catattaata	aatacctact	atgtgccagg	cactgcctac	2580
cactggaatc	tttactaag	acattgtttt	tactttgcat	ttctgccttt	acactatgaa	2640
agtagatgtt	ttggattcat	attcattcag	catacatttg	aatatgctgt	gttatgcata	2700
gtaagcctat	gataagcaag	tattctcatt	tagaatttgg	gaatattgat	tatacatgtg	2760
gacaaacaaa	ccataaatgc	aaactattta	tatgataaat	aactttggac	tgatggctgg	2820
gaggaaggac	cagctattga	tgggtaggaa	ctagcaagta	gcggactgtg	gcctgcatag	2880
accagaccca	tccgtagtga	tccagatgaa	acagccaccc	tcagacactt	ggataaaggg	2940
tccaccaagga	aaaaactcct	ggcctatcag	ctttgtatgt	acagttcagt	tactggaagt	3000
atttcctcaa	aagtgttttt	atggttgagg	tacacattcc	tacagcttta	cctgctgcca	3060
agtccctgtt	tcaagggaag	cagcaatgaa	ttacactgtt	cccgtagtca	aggacagtat	3120
atcttaccaa	gaactatacc	cacttaagga	ggtgctggat	gtcataaaga	tttgatcaa	3180
ccattatggg	tgttcagagg	agagattatt	tccagctcaa	gacccaggga	agaggacata	3240
ggatggatac	cacagtcata	gggaggattt	aacacaggac	atgtacacat	tagttagtgt	3300
ggtataaagt	ggaacagaaa	tgaatgagac	acaaagcctt	gaatgccaga	aatactagta	3360
gtcctgttgt	ggaaggatat	aaaactcaac	tgggagtgga	agagaaaggc	agcagtgaat	3420
ctaggagatg	tacagtaggt	tgaggtaaac	atatcctgaa	gactataatc	caaagattat	3480
ttttggtttg	aatttgtttt	ggtttgaatt	catggtatct	attttctttg	agtggatggt	3540
tggggagggt	ggcatgtaga	atgcattctt	accaaactcag	catgattttc	aagacagtac	3600
agagaaaaga	ctgctgagct	gatgtaggag	ctttggctgc	agtctctatg	gctttcagca	3660
agccgtttta	ccttactact	gcttcatgac	tgtggctaac	aaagtaggga	tagtacggag	3720
cacagaggat	ttttagggcg	gtgaaactat	taatactctc	tttgtatgat	actataatgg	3780
tgggtacatg	tcattataca	tttgcccaac	cccacagaat	acacagcacc	aagagtgaac	3840
cctaattgtga	actctggtct	ttgatgatgc	tatgtcagtg	tacgttcata	cgtgtaacaa	3900
gtgtaccact	ctagtgggtg	gaggggttat	tgataatagg	ggaggatgtg	catgtgtggg	3960
ggcaggaaagt	atatgggaaa	tctctctact	tctgctcaat	tttgctgtaa	acctaaaacc	4020
tctgtaaaaa	ataaagtcta	ttttttaaaa	agtggggatg	gtattacggc	aatataaaaat	4080
caaaaactct	tatgaacaaa	tcttttctcc	agatgtaaac	tgtcatatat	gcaccctcgt	4140
atgtgtatgt	ataattttca	ttcaaactgt	aaacaacttt	agaattggca	ccaaacatat	4200
aaacactgat	acattagact	atctcgaaca	ccttttactg	accactttga	aaacttgctt	4260
acctattaag	gttcattcat	agctgtgatg	ttctattttt	attttcaatg	tgggattatc	4320
ttctgtttcc	cccaggaggt	atattaccaa	attgtgtgat	ttgtttctgt	tggtttgatg	4380
caagatggaa	agccctacta	tgctcaaact	agaggtttta	tccaggacca	gtattgctgag	4440
aagagtgcag	cactgacgtg	gctcattcct	accctctcta	gccccagaga	ccaatttgat	4500
cccgctcct	atatcatagg	taagtttgac	aaatggcaca	ggtttttttt	taacttagtt	4560
aactctccaa	tattatgtaa	aagagtgtgt	tagtcagctt	gggctgtcag	gacaaaaatat	4620

cacagactga	gtggcttaaa	caacagaaaag	tcactttctc	acagttgtgg	aggctgaagt	4680
ccaacatcaa	ggtgctggca	acacggattt	ctggggaggc	ttttcttcct	ggcatataga	4740
tggtcacctt	cttgctgtgt	cctcacatgg	cctttcatgg	agtgagagct	ctttgggtga	4800
tcttcttata	aggacaccat	ttctgtcaga	tgaggggccc	acccttatgg	tttcatttraa	4860
ctttaattgc	ctccctaaag	gtctcatctc	caagtaccat	cacattgggg	attagggtt	4920
caacatataa	at ttggaggg	tggcgggggg	ggatgcaatt	cagtcataa	caaaaaaagc	4980
atgagtatta	ttaagtacaa	aaaaattaga	gagctttata	gaaaatatga	ggcatttttat	5040
gtagctggag	tgtgagtgt	atcagttatt	ttgagttaga	gcaatgtgca	tctactaaga	5100
agtggatgg	ataagatttt	tttggagtga	cccagggtta	aactgtacta	caagaatgta	5160
ttgctcagga	actaggttat	ttaggttact	tatttataca	aacctattca	aaaataattt	5220
aggaaagaac	tatcccagtt	atcccatact	tgcaaatctt	caatatgtgt	gcctctgcac	5280
gtacacatg	tcattcttagg	cctttatagt	ataaaggctg	atagttgaaa	tggcagctgc	5340
tgtgcttttg	ttaattttcaa	agctgccaaa	acagttgtga	gatagactca	caagaattta	5400
ctgattaata	caattttttaa	agtttttcaga	tttttacagt	tacttcagac	tttttatctt	5460
tctgcagtga	gcatgcatca	ttacttttgc	atcctgagaa	caagcataag	tgtgtttttg	5520
gagagaactc	cagggacaaa	taatatacca	ctgttattct	cacctatatg	tcaagtttga	5580
tacatttgca	aacaattcta	gccttctgtc	tataagtata	tagaattttt	atttacctta	5640
tctatggatc	aggatctcag	cagaggcagt	gatgtatcag	aatcaccttc	gggattcctc	5700
tactgcctcc	tcttttctaat	ccccagattc	tgatatgcac	ccttgtccta	cagcgaggca	5760
gcatggcatg	aggtcagaac	accagttctg	gagccagact	gtctagggtc	acagcctgcc	5820
atttaccggc	catgtgactt	tggcaagttt	cttagtctct	cttgccctac	ttcctcata	5880
tgtaaaatgg	gaataataat	agtgcctacc	tcagaagggt	gatgtgagga	atgaaggat	5940
tgatacatgt	aaacttagag	cagtgtgggt	acaaaataaa	catgatgcaa	gtgttcaatc	6000
actgtttttg	ggagaatgcc	atattcttta	agccgttaaa	gaagaaaaaa	gtattaagaa	6060
taatttcaaa	gtaatgcatg	tttcaagggc	taatgccagg	ttgctcccag	agtgggtctt	6120
cccagtgctt	agaaattttta	acatcttatg	aaaatgatat	atatgggtcaa	aaatgtattt	6180
aacctttccc	ttggctgcct	tccaggggcca	gaggaagatc	ttccaaggaa	gatggaatac	6240
ttggaatttg	tttgtcatgc	accttctgag	tatttcaagt	cacgggtcatc	accatttccc	6300
acagttccca	ccagaccaga	gaagggtcac	atatgactc	atgttggggc	tactctcca	6360
ataacaatta	aggaatcagt	tgccaaccat	ttgtagttca	caaattaaaa	ctgggtttcc	6420
aggcctgggtg	tgggtggctca	cgctgtagc	cccagctatt	gcaccactgc	tctccaagct	6480
gggcaatgga	gtcagattct	ccttcttaaa	aaaccacaaa	aaaactggat	ttccagttct	6540
ctaataattct	tagtaccaca	agatatgtca	taggtatctt	taaatgaaat	tcttagctgg	6600
aaaagtgact	aaaaagtttt	tctcctgcta	cctagtaata	aacaaatcat	tgtttattac	6660
tggtcactta	gaaaatttaa	agggataggg	ccaggcacag	tggcttatgc	tggttaattg	6720
agcactttta	gaggccgagg	caggcggtatc	acctgaggtc	gggaagtgga	tgcctgagg	6780
tcaggagttc	gagaccagcc	tggccaacat	ggcgaaaccc	cgctcgctact	aaaaatacaa	6840
aaattagcca	ggtgtgggtg	catgtgcctg	taatcccagc	tatttgggag	gctgaggcag	6900
gagaatcgcc	taaaccaggg	aggtggaggt	tgtagtgagc	caagattgca	ccgtgtgtct	6960
ccagcctggg	caacagagtg	agactcttgt	ctcgaaaaaa	aaaaaaaaaa	aaaaaggctg	7020
ggcacagtgg	ctcacgcctt	taatcccagc	actttgggag	gctgaggcag	atggatcgcc	7080
tgaggttggg	agttcgagac	cagcctggcc	agcatggtga	aacctgtct	ctactaaaaa	7140
tacaaaaatt	agccaggtgt	ggtggcgcac	acctgtagtc	ccagctactc	gggaggctga	7200
ggcaggagaa	ttggttgaac	ccaggaggcg	gaggttgtag	tgagcagaga	tcgtgccact	7260
gcactccagc	ctgggtggac	agagcaagac	tccgtctcaa	agaaacaaac	aaaaaattaa	7320
aagggataga	atataatgaa	atataattttg	aacttaaaat	atattctata	tgtgtatctt	7380
cctaggcaaa	agctgtaatt	tccagagaga	ccattaggaa	caggtagtat	ctatttttct	7440
ccattatttta	tttctagaaa	ctcataaaaat	ggattgtatt	tttctataag	aacaaaatat	7500
taattaaggt	atagatgact	gaccaagggc	ttaatcaaat	aaaatgacta	acagcatcta	7560
tcataaagcc	acacaagcct	tatgttctca	tctcaaaaat	gctgtgacag	ctttttggct	7620
gctttaacca	taagaaaaat	gattgggtga	tgatttttatt	agcccaggct	tttaaaaact	7680
ttcatctagg	ccacgtgcgg	tggctcatgc	ctgtaatccc	ggcactttgg	gaggcctgag	7740
tggatggatc	acttgaggtc	aggagttcag	gaccagcctg	gccaacatga	tggaaacctg	7800
tctctactaa	atatacaaaa	attagttggg	tgttatgggtg	catgcctgta	atcccagcta	7860
ctcgggaggc	tgaggcagga	gaattgcttg	aactcgggag	gtggagattg	cagtaagccg	7920
agatcgtgcc	actgcactcc	agcctgggtg	atagagcaag	actgtctcaa	aaaagaaaaa	7980
aaagaaaaaa	ttttaattta	atccttctgt	agaaacaggc	attcagaacc	attccattga	8040
tcttaataaa	gctgctcttt	actgtttcta	gtcaaaaatg	agacttcgat	caaaccataa	8100

gattttatac	tgcagatagt	cagcttcacc	aaagccgcag	aggaaacatg	tcgagatcag	8160
gcttctctgct	tgatagtctc	ttgactacca	ttaaaaacgaa	tattgggagg	tcagtgaagt	8220
catttggtagg	ccatttagcat	tgatatcttt	aaaacatcta	ccctaaacca	tctgctatgg	8280
acccataata	agaggcctgt	tgtatatgaa	attgtctaga	attcagggtgc	aggtccttgc	8340
cgggttaagta	agggagcaac	acgtaaaatg	ggagaggagt	gggggtgtact	cacttgccctc	8400
ctcttttgctc	ctgatttaac	cagcattttt	caaccctggg	aaaatttgca	gaatcctaagt	8460
tgattgtaaat	gatttttgagc	tgcagcagct	ttaactctta	ccctttttcc	acatagttaa	8520
gggtgtttgag	ttggaaagaa	acaactatag	gtagctacac	gtacataatt	atctctttat	8580
tcacaaaggg	tatagtaaaa	ttgattgtaa	ataactttct	aagtgccaat	attcaaaaact	8640
tttggtattaa	aatgtatttt	tcaccgtgca	tttacttttg	atgtatttat	ttcatttaaa	8700
caatttaaat	ggggctcttt	aacaaaaaat	ggtaatttaa	acaaaaacag	tatcgtactt	8760
agaatttgga	gtagaggccg	ggcacagtgg	ctcacgcctg	taatcccagc	actttgggaag	8820
gctgaggcag	gcggtatcac	tgagggtcagg	agttcgagac	cagcctgggtc	aacatgaaac	8880
cccgtctcta	ctaaaaatac	aaaaattagc	tgggcgtggg	ggcgtgcgcc	tataatccca	8940
gctagtctac	tcgggagggt	gaggcaggag	aatcgctgga	actcaggagg	cagagactgc	9000
agtgaagcca	gatcgcgcca	ctgcaactcca	gtctgggtga	cggcatgact	ccatctccaa	9060
aaaaaataaa	aaaagatttt	ggagttagat	catcataat	aagtaacaga	ttttaggaaa	9120
atcaaaaaat	ggctaataaa	atgaacacaa	tgtaaaaaat	ttattaaaaat	gtagactttt	9180
aaaaatctat	aaattgatca	tctgtttata	aattggcaga	tggttgtgta	ccatctttta	9240
aaataaagat	tgaatttcac	ccagtgtgat	ggttcccat	gcttatattt	ctcctgctga	9300
ggccggacct	gatatggccc	tggtctgtgt	tcccagcctt	gtttcctcat	taccactaaa	9360
atctttcccc	tgtatgcccg	cccaattttt	ctggctctga	gtccttgttc	atactgttct	9420
ctccaattct	accttccaaa	ggcctttctt	aacaccttcg	gattcctttt	ttgagaactt	9480
tccagattcc	catgcctttt	tggaaatcaat	ctctatccta	ttgtcatcac	atttaagttt	9540
ctacttccat	catcctcact	cctatccctt	tggtcctggg	atgacaggga	tgctgtgttt	9600
tattttactca	tctttgtaac	ttccacataa	cctaaccocg	gttcttgctt	atgggagatg	9660
ctgattgtag	gggtctgagtt	agatactgtt	aactaaaatg	cttggtgata	ttttagttat	9720
taattcatat	taactttggc	tgaaactttt	aaattctatt	gtgaatagtc	aagtaaaaaat	9780
tagattgtta	cattctgggt	tagtattaga	ttgtttttaa	gattgtttta	aacaagatgt	9840
ttttaagatg	agtttttaaat	agttctctta	acacaaataa	agcttaatat	gagtagttga	9900
aggaaattat	cccaaaccat	tccagtctct	ggctgtgaaa	ggcttttcca	ggcctaataa	9960
gttttccact	tcagccgtaa	gtaggtgaaa	tcaaataaac	aatagaggga	aatgtattta	10020
tttgctttat	acacatgcat	gtgtgttgtg	tctacatata	aacattgcac	acgcttagaa	10080
tgaagtttct	gtcatgcccc	gaaaagggag	aggcattttt	gtggattttg	tctggctgcc	10140
gtgggtgtgt	ttgaagaact	gtgctgttta	cttcatacca	gggtgtgtgag	ccataccttt	10200
ggtaggaggg	tatacctcct	acaccaaga	aatataagcc	aggagaaggt	ctgtgccaa	10260
agaagggaacc	caaatagccc	acaagagggt	ggccattaat	tattgggtca	gatgcataaa	10320
tgacacagtaa	tttattttaag	cacctcttaa	tggtgaccca	caagggaagat	tgctcgtagt	10380
agcggaagag	ttcacaataa	ataagagaaa	aaagcagaat	gtagaactgt	atgatagcaa	10440
ttctgcaaac	aagaagcacc	ttttataaaa	gatggaagga	gcccaggcac	agtagctcat	10500
gcctgtaatc	ccagcacttt	aagaggctga	gggtggaggat	cacttgagct	gcagtgaacc	10560
atgattgtgc	caccactcca	gcctgggtga	tagaagttag	accttctctc	aaaaaaaaaa	10620
aaaaaaaaaa	aaagacggaa	attcctccag	aatttttaaca	tgtcaacaga	ggttttctgc	10680
agctactttt	ttcagcttta	tacttcgcag	tattttccaa	attttctcta	acaagcagta	10740
ttttccaaat	tttttacaat	aagcacacac	acacacacac	gtttgtttgc	ataagtcccc	10800
aactgggtgt	gaacaaccgc	tggtttttta	tctatacata	tctagaatat	tttataaata	10860
gtagttctta	aacccttgaa	agggagttaa	tgaccagctg	agaaaaataaa	gtcagtgatt	10920
tcattatttt	cctatatatt	catcatgatt	ctaggaaaga	acttgaggag	gacttccttc	10980
agcttcagcc	actcctgggc	caggcgcatg	cttagctctg	tggtaaaggt	caccagcttc	11040
ttctgcaggg	tgccctgtatc	atctgaattg	gagggtttggc	gagggtaaga	gactgatgta	11100
ggttcaagtt	tttctttctt	gtcctccact	tgaaatctgt	cttcccttcc	agactgcctt	11160
cgctcgtgac	ttaaggcccc	aacaccaaac	acagaagcaa	cagccttaca	cagagtgttc	11220
agcaagctcc	aacaattgtg	taaggtaaag	tttcttttat	agattccttt	tctatatcgc	11280
tcctagtggg	tctgtttctc	tgatcgaatt	ctggctgata	acagttgctg	agactctgaa	11340
agagaaggca	aggaaactact	gtttctcatt	ataaactgtt	tagaattatt	tggccatctt	11400
tttgctatga	atatgtagtg	ctttgatata	tttttttaaat	caaaaagtaa	tgaaagagat	11460
cacataggga	aagatagatt	ggattatttt	taaagtttat	atactaaatt	gaaaagcaaa	11520
gaataaaatg	ggagaaacag	ctccctcatg	tggctgttgg	cagggaagctt	ccattcctct	11580

ctgtgggcct	ccacaggttt	gtcacagca	aatggctcgt	gacagaaaga	cgcaagggca	11640
gttgacacca	agatggaagc	caccatcttt	tctataacct	aatctgaaag	aaggacata	11700
ccagcacttc	tgccatatgc	tgttggtgca	cacagaccac	ctctggtaca	gtgtgaacac	11760
aggaccacac	aagggcggtga	attccaaggg	cagagaccac	tagggaccac	ctcagaggca	11820
cagaggggaca	ccctatccag	ctggtggcca	atgtaaatta	acatagcttt	ttagaatagc	11880
aatatgtatc	tataatctta	aaagtattaa	aagtacttct	tgatccagta	atttcatttc	11940
taagaatcca	tgctaagagg	atttaaaatg	tggacccaaa	aatgggtata	aaaagaagtt	12000
gttaacagta	tttaaagttg	tgaaaaacca	gaaacaatct	aaaggtccaa	caataggaaa	12060
atgaattttg	atatttttct	aatagaatgt	tatgctgtca	tcagaaatac	catttacaaa	12120
taatttttaa	taacgcaaaa	aaaagtttat	aaaatgttta	gtgtaaaaac	tggacacaa	12180
tacataatga	ttctgatttt	gtaaaaaaa	aaaaacaaaa	cacacacata	tacacatgca	12240
tacatatgca	tataaagaaa	actggaacaa	acaaaataac	aagcatagtt	ggaattacag	12300
tcattttaat	attctttatg	cttttaaaaa	ttttgaagtt	tgtattacta	gcacccacta	12360
cttacgtagt	caggaaaaaa	atacaacttt	aaaatagata	tttaggtcca	aagatggtaa	12420
tctaaatggt	gttacaggct	gaatgtgtgc	ctgatcccca	tgcccccaag	tcatatgtta	12480
aagccctggc	ccccaggcca	atggtattag	gggagtaggg	cctttgggag	gtaatcagat	12540
ttctacagag	tcatgagggg	ggagcccgcga	tagtggaaat	agtgtccttt	taggaagagg	12600
agaacagacc	aaagccttcc	tttctctcct	cactatgtta	gaagacagcc	agaagtggtc	12660
cacagccagg	aagagagctc	tcaccagaac	ccaaatctgc	tagcaccttg	ctcttggggt	12720
ctcagcatcc	agaactgtga	gaaatgaatg	tgtgtgtgtt	aaaccactca	ggctacggta	12780
ttttgttgca	gcagcccaag	ctgacagaga	tagaaacaac	acaaggaccc	atcagcagac	12840
gaatggatga	tcaaaacgtg	gtgaggtcgt	gcagtgggat	attattcagc	cgtagaagga	12900
atgaaatttc	gatacatgct	ataatgatga	accttgaaaa	catgttaatg	gaaataagcc	12960
aaacttaaaa	ggacaaatat	tgtataattc	cacttatatg	agttagttac	ctagaatagg	13020
caaattatgt	catagataca	gaacattaga	ggttaccagg	gttgtgggaa	gaggggtatt	13080
gtgggtacaa	attttcgggt	tggagtgtat	ttgaaaaaat	tctggaaatg	ggtagtgcac	13140
gtagtcaaca	tgatgaatgt	acttaatgac	actaaattgt	acacttaaaa	atgggttaata	13200
ctgggctggc	gcagtggctc	atggctgtta	atcccagaac	tttgggaggg	caagacaggc	13260
ggatcatgag	gtcaggagat	tgagaccatt	ctggctaaac	tgggtgaaac	ctgtctctac	13320
taaaaataaa	aaacaaataa	aaaaaaaatt	agccgggcat	gggtggcaggc	acctgtgctc	13380
ccagctactc	gggaggctga	ggcaggagaa	tgggtgtgacc	tgggagtcgg	agcttgagct	13440
gagctgagat	cgcgccactg	cactccagcc	tgggcaacag	agccagattc	cgtctcaaaa	13500
aaaaaaaaaa	aaaggttgat	acctgggtgc	ggtgggtcat	gcctgtaatt	tcagcacttt	13560
gggaggccaa	ggcaggcaga	tcagttgagg	tcaagagtta	aggaccagcc	tggccaacct	13620
ggcgaaaccc	cactctctatt	aaaaatacga	aaattgtctg	agtgtggtgg	tgggtgcctg	13680
tagtcccagc	tgctgggagg	atgaggccta	ggaattgctt	gaaccaggga	ggcagaggtt	13740
gcagtgagtt	gagattgcgc	cactgcactc	cagcctgggg	gacagagcga	gacttagtct	13800
caaaaaaaaa	gttaaaattg	taagttttgt	tatgcatatt	ttaccataat	ctttaaaaaa	13860
tagatatata	ggagataaaag	tcaacagaat	ttaataacca	gttgtaaata	gagactgagt	13920
gaggaggatg	aattaaggaa	gacattgagt	acaacttttt	ggtaggtgaa	aaactcttaa	13980
aaaaatacgt	gggcaaagat	cctacttgat	tcttataaatt	taaaaatctc	ccagttagta	14040
aacaaggcta	ggtggagatt	tgcatgtgat	gtgaggtgtg	tgttctgttt	tgtaatgtga	14100
ggactgtgag	ccatctcctg	gacttgaata	tccattagat	aattgaaaat	acggatttga	14160
gaactcagga	gacgtgcaat	gcagtaacaa	aactctgcac	ctagttgatt	tctgtctcct	14220
aatttaatgc	ttttatggga	caaactgtta	ggcaggtggg	caagatggac	agccatattt	14280
ttgtgggttt	ctggcctgtg	ggccagcctc	agtgtcctct	ctgaggtcat	gtccaaactt	14340
agaacacatt	caggcctacc	acagtcaagg	ctccctttct	caactctagt	cctctgcaca	14400
aatatccgaa	gcctagaaat	aataatcatc	tgtcctttgt	tcttgcatat	tgaagccta	14460
ggaaagggcc	ttgggaatta	agaagaatgg	aaaaactggt	ctaactgctg	catgcttcag	14520
cttgcaaggg	aatcactgaa	atggggacag	gccataaaaag	gacaaccaga	agagtggtct	14580
cagcaaaggc	atcgtttttc	agagcaagct	agagaatcct	gccagcgtcc	tcaggcaggg	14640
cccctgggca	cagaggttag	gcaagggagt	gtcccagcat	gttgatgccc	tgaacatcag	14700
aataatgccca	tagaggagct	tccaaagagt	tcatttcagg	ttttgtgaag	cgaacatttc	14760
taggcaataa	aaattttgatt	ttgtgaataa	agcttgtttc	ttcaactcca	gtgcagattc	14820
tcatagattg	atagtggctt	gtgatccaga	taaagaaaac	aatttttcaa	agattcatat	14880
tctttgtaga	tgtacggatt	tagagaccat	ctaacttaac	tccctcattc	tacagatagg	14940
aaaaatgagg	cctaaagaag	ttaagaaaaat	accatggaaa	tgtcactgct	gaactgcat	15000
acgtaggatc	cgaaagaaat	tgggtaaatg	ctactgtgag	aaatacagta	ctaggtccaa	15060

agaatcta	ataa	aatctaa	ttatttcta	agcatccctg	cacatggctg	15120
aacttacata	gtttcatttt	ctttcttttt	tgttgaagaa	gaggcaattg	gctgggtgca	15180
gtggctcatg	cctgtaatcc	tggcactttg	agaggccgag	gcgggtggat	cacctgaggt	15240
caggagtttg	agaccagcct	ggccaacatg	gtgaaacccc	atctctacta	aaaatacaaa	15300
aattagctgg	ctgtggtggc	cgctgacctg	aatcccagct	actccagagg	ctgaggcagg	15360
agaattactt	gaatctggga	ggtggagggt	gcagtgaacc	aagatcacgc	cattgcactc	15420
tagcctggat	gacaagaggg	aaactccatc	tcaaaaaaaa	aaagaaaaaa	agcaatcact	15480
aacctgtgtt	gtttattaaa	catgacagac	tggcatgaag	taattaccaa	actgtaaaca	15540
aaaaagctac	aatctgccag	gcattggtgg	tcatgcctgt	aatcccccac	cttgggaggc	15600
cagggtgggg	gatcacctga	ggcctggagt	tcaagactag	cctgggtcaac	atgggtgaaac	15660
ctcgtctcta	ctaaaaatc	aaaaattagc	ccggcgtggt	ggcacatccc	tgtaatccca	15720
gttactcagg	aggctgaggc	aggagaatca	cttgaacctg	ggcagtgggg	aggttgcaat	15780
gagccaagat	cgcaccgttg	tactccagtc	tgggcccaga	gagtgaactg	cgggtctcaa	15840
aaaaagaaaa	aagaaaagct	acaaccttaa	tctcaacttc	tcataacatc	atctctactt	15900
ctgattagaa	gagtggaaat	ggggagggtt	attacaaaaa	gactgttata	ccttacacac	15960
ttctcccat	gaatagttaa	ggtgtgagtg	aaaaagacag	caattttatt	ttttttttga	16020
aacaggttct	tgcatgttca	cccgggctgg	agtgcactgt	tgtgatcact	gctcactgca	16080
gcctccacct	cccaggctca	agtgatcctc	ctacctcagc	ctcctgagta	gctgggacca	16140
cagttgtgca	ctaccatgcc	cagctatttt	tttttaagag	atgggggtctc	actatattgc	16200
ttaggctagt	tctcaaaactc	ctggcctcaa	gcagtcctcc	gaccttgccc	tcccaaaggg	16260
ttgtgattac	aggcataagc	caccacaccc	agccagcagt	tttagaataa	aggggtgaagg	16320
tgctgttggg	gaaatataat	ttaaaaaaca	aaatcttctc	tcaaccaga	aatcctctcc	16380
atgaaggcag	tagagaaaaga	taagctttat	tattgaataa	aaattaaatg	agaatgtgat	16440
gcacatcaca	ggcactttgc	taagagatca	caaagacaga	aggaaaatttc	accattttgt	16500
acagccaagc	aggtacagcc	cattacatgt	atgttttcga	gataaatagt	cctcaactaa	16560
gagaacttga	cagcaccact	ggtcacacag	ttcattctaa	ctttacctga	taattgatgt	16620
gaccacttgt	gttatctaa	atatcaactt	ttcgggggtg	ggggagtggt	gaaacaggag	16680
ttacttttat	agcttgggtg	aaggtactca	ttaagattag	gctgttacct	tcccacagaa	16740
actggaagat	aggtatgcta	tctggtaatg	tttacatttc	ccagatcctt	gagaaagaca	16800
ttcctaggct	ataaagctga	caaaaggctg	attcagtttt	taaatatata	tatctgtata	16860
tgtatttca						16869

<210> 65

<211> 15000

<212> ADN

<213> Homo sapiens

<400> 65

gatctcttga	tcccaggagg	tcaaggctgc	aatgagctaa	gatcaagcca	ctgcattcca	60
gcctgagtga	tagtgggaga	ccttgtcttt	aaaacacaca	cacacacaca	cacacacacg	120
agggcctttg	accactcttg	agtagaagac	tcgagaagaa	caaagtagaa	ggccagagaa	180
gaacaaagtt	acttgaaaga	tctcttatta	aagagaatgt	acaagctatg	aaaaaaaaaa	240
aacacacaca	cacacacaaa	cctcatctgg	aatgaaaaaa	acataatgca	tttggtttct	300
ggttccttag	gctgttatgg	aacaaccaa	gaacattatt	ttggtttctg	aggtcagaac	360
tattttattc	ccctcaagca	cactatgctt	atggtttgag	ggagaatgag	aaataggaaa	420
ctaggaacag	gctgaaatgg	tctaattctt	accatcta	tctgcagtgt	cttattctca	480
ttctaaaaga	gaatggttat	attcgtgtgt	ctagcataaa	aagtaatgat	aaaaataaaa	540
gatcccgtat	taccagacaa	taatccccta	gactgtttta	atgcttggtt	gagtatttgc	600
ttatgatctc	agacttttaa	agatgggtct	cccctatggt	gaagcttggt	aattatgtag	660
gcacatttaa	tgtctgttta	cttatcaaaa	ttttatcatt	gttagttgta	ttactacttg	720
acagttcca	ttatttaatt	gaaaagattg	gttaacattt	tatagtcaaa	gtaattgttt	780
cctgtgtttt	ttcctgttta	ggttattgga	gttagtgagta	aagaatacat	accaaagggc	840
acacgttttg	gaccccta	aggtgaaatc	tacaccaatg	acacagttcc	taagaacgcc	900
aacaggaaat	atttttggag	ggtaagtaag	ggaaatttct	tcagacccat	taaattgttag	960
gaaaaaatgg	agctaaaaga	gctgggtggc	tcacctttct	catcctgtgc	tgagaaatgc	1020
tggggctcac	ccataagtat	ccagcatccc	catggacaca	gggaattctg	aacaaatgtg	1080
atgaaaccga	tgaaatgtct	ggcctgtagg	tggttagtga	tggagatacg	ggctatatgt	1140

gaatcttgat	ttttgcaatt	cattagagct	ttgtaatgaa	aggaaacagt	ttgttgcttg	1200
ctttaaggat	aggttcattt	gcattttctcc	gcaaggaagt	agtaatgagt	taccaagcc	1260
tagatttcac	ccctttttga	tttcttgctg	acttaacttt	aattgaatgg	aagagttatc	1320
acaaatgaat	tatctttttg	gttttttttt	ttttgagatg	gagtcctcac	ctgtcaccag	1380
gctggagtg	aatggcatga	tctcggtcca	ctgcaacctc	cgctcccag	gttcaagcaa	1440
ttgtcctgcc	tcagcctccc	gagtagctgg	gactaagggtg	cgccgccaca	tgcccagtta	1500
atttttgtat	ttttagtaga	gacggggttc	cactatgttg	gccatgatgg	tctcgatctc	1560
tggacctcgt	gatccgcccc	ccttggcctc	ccaaagtgtc	ggaattacag	gcaagagcca	1620
ccgcgcccag	ccaggaatga	caaatagaatt	accttataag	taaatgccat	taaggaagga	1680
tagctggaag	atgggttgag	gggaatggag	gaccacagaa	ctagtcctat	ttaaatacat	1740
gtgcatggta	aaatgattcc	atttgacaat	aggtttaatta	tctcatagca	taaggaaaat	1800
gcttaacagt	catatgcaag	atgataagct	ttcctatagc	atccaaccaa	aagatctagc	1860
cagtacaatt	tcctttgcta	tattagggtt	agaaaggccc	ccagaggtga	accaattaga	1920
tggaaacctt	gaataaaaaca	ctggattagc	agtgaacaga	aaaaagtcag	attgctttcc	1980
ttcttcccat	agatgtctca	gggatattta	gtttcctcag	aagataaaga	atttagtaag	2040
cgtttttttg	tgcatactta	catgaaatgt	acattatttg	aattctttta	aaagaaacag	2100
ctgcatgata	acaaaaattg	tgttatgctt	gcttttagctg	gtattttttgc	ctagaacgat	2160
tatatcggtc	ggacaagaag	ctattcctaa	gaaacaatat	ttttaatcca	ggaagttttt	2220
cattttttaga	aatttatctt	actattttcc	aagcaaaaaga	gggtagttac	agattcacta	2280
agaatcatgt	gtcacaattt	tttattttaat	aattattcct	ccttaaaaata	tattaatcac	2340
ctgacttaca	atgggtggaac	catgagtgca	tttttgccct	tattgtcaat	aacgtcttct	2400
cagaagttag	ccacaaaggt	gcatagttct	tggagttaaa	ggtctgaatt	aagacaatcc	2460
agcataagtc	tcattaatgt	gtgattattt	tgagaaaagg	caagaagtac	ctaagaatct	2520
ccccctcact	gtccagttcc	ctgtttcatt	taaagattca	ctgtaagtaa	ctgaaaggct	2580
ttccttgggg	ggattttatt	gaatcagttc	ttcacatgca	aaggatattg	tagaacatct	2640
cgtttttgct	ggcaggaata	tgaacatctg	ttgtgaggaa	agaaaaagtc	tcatgcaaat	2700
tacactgcca	aagaagggat	gttcaagttg	agaaaccagt	gacattttct	gtaactgtac	2760
tatgaatcag	cgcattttaa	tcttctagat	aatatatgga	agtgcaggaa	ggtggtagga	2820
aacggtgttc	attttacata	tgcgttattt	tattctgtgt	gagtgacttc	attggcaccga	2880
cattgctgtt	tttaaatgag	gatacagtaa	attgcagtcc	gaggaaggct	aactggaatc	2940
aacatacccg	tagctttaga	aagcagtttc	cgcaccagcg	aagagtacaa	gagcgatgga	3000
accccatgtt	cctggaagtt	tgcacatcag	agtaaaacaaa	cttgaaaacc	cctcttgata	3060
gcagaattca	cccagccttg	ttccattttc	tcttaacaaa	acacaccgca	aaagctctca	3120
caagctgctt	tgatgaagcc	acatgtattt	cccccttcac	aattttacagg	aagttactct	3180
taaaagaaa	tgattctggt	gtttaccgcc	tgtgtttaaag	ggacagagtc	cctttttatt	3240
tctgataacg	tttgagcgaa	atacagaaac	tatctgtaga	ctagcatagt	cggtacgtga	3300
gtaaggaaaa	gcaataacct	gctgtccggt	gagcacaaaa	ttcctgctac	gaacagtgc	3360
ttactgctgc	ttggagactg	caagtcgcag	atcacactag	gtattgactg	attgtataag	3420
gaaatttctt	aaagtctaaa	gtaaagggtg	tacctcctaa	aaagagggga	agagagaaaa	3480
ctttgtgtgg	aaggataagg	agtgtgttta	tagtttcagt	aagagtgtac	gttttaattt	3540
ttcttcttcc	tctgcctctt	tgccaagtag	cctgagtgca	tctgttatcc	agaagttagta	3600
ttactctagg	acaaacttca	aattcttcat	tctgcgttgc	ctttaaggaa	caacatactt	3660
tcttctgtgt	ctttttccaa	aaacacacgc	ctatggctct	gtgtgtgggtg	ttttagccag	3720
cctcctccca	gataaggggt	tcccttccct	cctttgcatt	gaaaggaaa	tgcaagtctg	3780
gacatgttta	tcaagaggaa	aagtgacttc	tcagtaatat	actgtcaaat	tcgggctgct	3840
gcccagagtgt	tcgctttgtt	atggcaggtg	aagttcacct	ttgccccacc	cagtgtttcc	3900
acaaaaaggc	aaggttccaa	gtattcatat	gaacaagtgt	tacttttagga	cttgagggt	3960
tgggggtgga	ggatgtttgc	atagttgaag	ccttgggcgg	gggtgtagga	aacggcgagt	4020
acagaggcca	tagaaaaagc	taagactcag	tttgacgtcg	tcagccggct	tggtcttcta	4080
cccagtgact	caaagcacta	aaagtacgca	taatcggaac	tgaagttagt	agcatcgccc	4140
atgtgccatt	cactgcagta	gcaaaagtag	tactctgtgg	tgggttaatc	ggttttaggc	4200
agctccttaa	atgaacattt	gtgtttcatt	tttctgttat	tttcccgaac	atgaaaagac	4260
gataaaaactg	aaatggaaaa	ggtaactgac	aaaagtgtgc	cttacctgtt	tccgccctga	4320
tttctgctga	ttcaagacta	ttctggctaa	actgattgga	ttctttttct	aactaggcag	4380
taggggatca	gaaatcacac	acggtaccgg	ctgtgtttat	tctgagagg	gctggggagc	4440
tttgggtctg	acttcctttt	acatgcctgt	cttctctttt	ggacagatct	attccagagg	4500
ggagcttcac	cacttcattg	acggctttta	tgaagagaaa	agcaactgga	tcgcgtatgt	4560
gaatccagca	cactctcccc	gggagcaaaa	cctggctgcg	tgtcagaacg	ggatgaacat	4620

ctactttctac	accattaagc	ccatccctgc	caaccaggaa	cttcttgtgt	ggatttgtcg	4680
ggactttgca	gaaaggcttc	actaccctta	tcccgagag	ctgacaatga	tgaatctcag	4740
taagtggatt	acagaacaaa	aaaataaaaa	atgccagtaa	tgtcggttct	gcccccttga	4800
actaataaca	tgttgtttta	ttatacggct	ttgtcatgtg	ttggatgaag	taggtggctt	4860
aagctaggga	ctaggaagag	gaaaaacatt	ttttgagtc	ctattaacta	ttaggaaact	4920
tgatca-tta	aaagtatata	tatatatgag	gagctacott	gagttttgaa	ttcaggatgt	4980
tacaggaaga	aatatatgtc	caatttcta-	ttatccaaaa	gcagttggga	gaattacagg	5040
gatttggcca	gacatgctgc	gtatgcaagg	tatagccctc	atctgtggta	ctttggcagg	5100
gcttagactg	catcaaaaata	tttatagatg	tacatttgag	tgtacagtta	ggatctgatg	5160
tggaaacattg	taagatcatt	gctagaaaaa	ctttgtcata	atttttcaat	attattctaa	5220
gtgaataaacc	gtaaagattt	tacatcttag	cttccttcc	tacagtaaaa	aaactatctg	5280
atctcttgat	cagtattata	gtagccacct	atcactttat	cttaacaaat	tctcaattcc	5340
ttaggtttat	gtgcttttac	ttcttttatt	tgattaaaaat	tgctgtcatg	acctctctct	5400
gcagagggct	gcatcatttt	ggtcattctc	aagtgatctc	tttgagcaat	ttaagaattg	5460
ccataagatt	ctaacctctg	ctgtaactat	ggttgtgtgt	tcttggttag	accactaaat	5520
cttattagca	gttttaaaaa	ttattccttt	tggtttagaa	gttaagacta	aatgctgaag	5580
tttttgaac	ttttggtttt	gatatcattt	caaacttaag	aaaacatttg	aagaaaaagga	5640
caaagaattt	ccacttacct	tttaccagg	tttaccagtt	attgataagt	atatccattt	5700
gctttaccag	aaggctaact	tgttttagtt	ctcattttca	cctttgagac	atttggataa	5760
aatatcaatg	ttaacataaa	ttggaatttt	gactttgatt	ttaggaccaa	tgaacaagcc	5820
aagtacttac	cctagtcata	tataatccaa	ctgtatgggt	atttggtatt	cattccacac	5880
ttcattttac	ttgatctccc	ttaaagattg	aagatttgtg	ttgcagtttt	tctgaaaatc	5940
tggggctata	aaagcatcag	gacctccccc	gtaggggagg	togtgtgttt	gggttcctta	6000
cacaacaggt	taccttgag	cttcaggaaa	agaactggct	ctcagttccc	cagttccagc	6060
ttaatgggtc	taattaggtc	ctgaccaaaa	aggtggcagt	tcttttccct	catgtctctt	6120
cagcgctccc	cgagactctg	gagactctgt	catatcccta	gggctgagcc	tcccagggaac	6180
cattcggctg	ttgtggcatc	tgtgtatgcc	atgccagtg	ctgaggacct	agtaacaaac	6240
gacaaatgca	caggcacagt	ggcatttttt	tggaaactcg	attccagctg	tgcgtctcag	6300
aagaagcgca	cagctccctc	ctggctttct	taacatagtg	agccacttcc	acttaagggt	6360
ctccttaacat	tctttgagtt	taatcattca	ttgattcaga	ggaaagtctt	ttgattttttg	6420
cttttcttta	aacagttcat	ttgaggtgac	ctaccccagt	gactttgcac	caaccaccaa	6480
gaaacttttt	tgcattgctt	ccgcaccctg	tgccaatcaa	gggaagggtt	taaaggcctg	6540
gcgtttttat	tcttcaaaga	aagggttttg	acagtatttt	aagggttcaag	tgcttctact	6600
ttgtgttcag	aagcaactgt	catatatact	gtgaaatgac	acctttttat	tatccctttt	6660
tatttatgca	gtatgtcccc	ttttattttg	gcagaatttt	ttctaaatgg	ttgttttaaca	6720
ttttcaagca	catttccattg	tccaatatcc	atagtaaaga	atgagagtta	acaataacca	6780
gtcacattaa	aacaagattc	ctgctgccag	ttgtgaaacc	ggttgtctta	ggcgtggcag	6840
ctgatgattg	agactgtgat	caggaaaatt	tccactattt	catcaggcct	aataggtaga	6900
ttgtgtctcc	aaatgaactg	tgttgggttt	ccatgcttaa	agcacaaatag	aggtgggtgca	6960
agaatctcca	tgagggtcta	aatggcagtg	atggttcagg	cggtagagtt	tgagaagaa	7020
gggatttgaa	acaaaccaaa	ggaaagaaaa	gtaagtagcc	agaaatcaca	aaatggcatt	7080
tttctaaaaa	caaaggaaaa	ggaataaaaag	aactaataag	tttgaaaccc	ctacccctcc	7140
caaatttggc	agggggggag	gtattttttt	tctatctatc	taactaaccc	atctagaaaa	7200
cagttgacca	aattatagac	ttctaaatgt	taatctgctt	tctcagtttc	agttgaaaaag	7260
agactttgtt	ttgcctactg	cagaacttct	aggttctttc	ttatagtctt	ggggttctta	7320
ttatagatcg	aaaatgtgag	tccgcataat	taagccattc	ggagtcttca	gaagcagttc	7380
actcttgaaa	tgactccgtc	cgccctacagc	catttaagat	ttcagaacaa	aaacagatct	7440
tgattttctt	tttcatgtta	actcaagctg	ttgctgagtg	ggagagtcag	aaatgacacc	7500
agctccactg	attactcagc	tgctgaagga	tgatttttta	aaatgcacct	ttactgtata	7560
tggacttcct	aatttccacc	tgtagagcat	cttagggagg	ctaacatgtc	actctggatg	7620
ttctttttaga	ataagatgca	aatctatttt	tctgaaggca	ttagagatag	caaacattta	7680
ttgtgagttt	actatatact	aggcactgtg	ctaagtgttt	tgcatagaaa	gtttaaaaat	7740
ctggcttttt	tgttggccca	atcataagtt	tcatatcagt	tcaacattca	aattatatta	7800
aggtacttaa	gaagaatccc	tggctaaatg	tgagggggcag	tgccacagat	ggactgaaac	7860
tttatgctta	ttgcacattt	atgctattat	tatttgttga	attatagaac	caagggagtg	7920
tgggaagccac	tggaaaaaat	atgagactta	gatacataat	ttgagtaaaa	atggctcaaa	7980
gtcatgaggg	taaagttttt	tgtatttcca	ttttattcga	gcggcatcgt	ttttaaaaat	8040
cattatgaat	ttgaccctat	atagatgttt	ccaaataaatt	ctttttcacc	ttcataaaat	8100

tccttccgtg	ggctgtgaga	tgccttgcc	atcagttttc	aagcttagt	gtctttctca	8160
tcctttacca	tttttagctt	aaaaaacaaa	agtgacaatt	agaacttcct	gcctgctggg	8220
cctcactgaa	agaccgatat	tggcctgata	aggagatatt	tattttgttt	tagtggcttc	8280
agaaatccct	ctccctcagc	aagctttcca	tcacggcccc	cccgtcagca	tcttccctga	8340
tagcgttctt	ctctgtgttt	attctggggc	ttcaggctcg	cccaggagga	actgataacc	8400
gctggcagga	gataacattc	tctaaggggc	tctcaaattg	gaatcgaatc	cctcaagcca	8460
gtcagccctag	agaatacatt	taaaagggtt	agttctggag	tttcacagag	ttcatttcta	8520
gacctatcag	atagcaagtg	tggagttctt	tctcaactaa	attcaagcag	agacattttt	8580
tagacgatga	aggatatttg	cacaaaggct	tcagcatgat	cccccaaac	tgctgcctct	8640
gaaggcatct	ccacacattg	acagccaatg	ccttcagtcg	gttcctaggg	caggtgtcct	8700
ggcttgagt	actgtcctcc	aataatcaga	gctcaacta	aacatcgtat	gttttacttt	8760
tggtttccag	gcaaggctga	gcagggaatt	ttcagttttc	cctgcccaga	tgggtgtttt	8820
ttcctgaagg	catcatttat	tgtgtagcga	ggagacaggg	ctggctgtgg	cagggatagt	8880
ctagaactgt	cctcattgct	gctgttctta	aatagtatct	ttaccaagta	ataacgtgcc	8940
gtctttggga	ataagtgcct	tcctcttagc	ctgttctgtt	ttcttgggtg	cgctaagtaa	9000
ttgaactggc	tcagggaagta	cctattgtgg	tttggcagag	gtgactgtca	cgcttctgtga	9060
ctccaggggc	cagcactgct	gggatcctgg	ctagaccaga	cagagccttg	gtgaagtgtc	9120
taggtctgtc	gcacatcgcg	aggaaggctg	tattcacttc	gctaagctcc	ttggcatagg	9180
cagtttgaac	agggctttat	caaattcgta	ttcaacaaga	gtagaagcga	aaattgatga	9240
ctgtgtatta	cttgaaatga	gtcttaatct	ttcacattta	gttctcaggg	tatgctgatt	9300
tccttttaggt	aaaccatgaa	catcagaaag	acttttatta	acctatgaca	gggtccccac	9360
cccagtatatt	ttccactcca	ttaaaatgga	agtttttttt	ttttttttct	tttttgagac	9420
agagttttgc	tcttgttgcc	cagtctggag	tgcaatggca	caatctcggc	tcaccacaac	9480
ctccacctcc	cagattcaag	cgattcttct	gcctcagcct	ccaagtagc	tggtgattaca	9540
ggtgtgcgcc	accacgcccc	gctaattttg	tatttttagt	agagatgggg	tttctccatg	9600
ttggtcaggg	tggtctcgaa	cttcgcacct	caggtgatcc	gcccacctcg	gcctcccaaa	9660
gtgctgggat	tacaggcaag	agccactgca	tccagcttag	gctatcttac	tccagcctaa	9720
acagcaattt	tctatcataa	ggtctgtact	aatgaaaaa	gaatcaccca	aggtctgtgt	9780
ttgttctgtc	tgtgtgcga	ttgtccgcct	tttgcctgag	aggaaacgga	actgcacttt	9840
tgagtgatgc	gccagagcgc	ttctagaatg	agagtgcgtt	ggaagccaga	tatgtggcga	9900
ttgtgtcgcc	agctgttact	caggttttct	caagaaggag	gagcaacttt	ggcagttttg	9960
cttcagttct	ctctagccct	ctgtgtaatc	gccccttttt	ctttattttc	gcacaaacac	10020
agagcagctc	aaagcaaccg	agcactgaga	aaaatgaact	ctgccccaa	aatgtcccaa	10080
agagagagta	cagcgtgaaa	gaaatcctaa	aattggactc	caacccctcc	aaaggaaaag	10140
acctctaccg	ttctaacatt	tcacccctca	catcagaaaa	ggacctcgat	gacttttaga	10200
gactgtggag	ccccgaaatg	cctttctacc	ctcgggtcgt	ttaccccatc	cggtccctcc	10260
tgccagaaga	ctttttgaaa	gcttccctgg	cctacgggat	cgagagaccc	acgtacatca	10320
ctcgctcccc	cattccatcc	tcaccactc	caagcccctc	tgcaagaagc	agccccgacc	10380
aaagcctcaa	gagctccagc	cctcacagca	gcccgtggaa	tacggtgtcc	cctgtggggc	10440
ccggtctca	agagcaccgg	gactcctacg	cttacttgaa	cgcgctctac	ggcacggaag	10500
gtttgggctc	ctaccctggc	tacgcacccc	tgcccacact	cccgccagct	ttcatccctt	10560
cgtacaacgc	tcaactacccc	aagttcctct	tgccccctta	cggcataaat	tgtaatggcc	10620
tgagcgctgt	gagcagcatg	aatggcatca	acaacttttg	cctcttcccg	aggtctgtgc	10680
ctgtctacag	caatctcctc	ggtgggggca	gcctgcccc	ccccatgtc	aacccacttt	10740
ctctcccag	ctcgctgccc	tcagatggag	cccgagggtt	gctccagccg	gagcatccca	10800
gggaggtgct	tgtcccggcg	ccccacagt	ccttctcctt	taccggggcc	gcccgcagca	10860
tgaaggacaa	ggcctgtagc	cccacaagcg	ggtctccac	ggcggaaca	gcggccacgg	10920
cagaacatgt	ggtgcagccc	aaagctacct	cagcagcgat	ggcagccccc	agcagcgacg	10980
aagccatgaa	tctcattaaa	aacaaaagaa	acatgaccgg	ctacaagacc	cttccctacc	11040
cgctgaagaa	gcagaacggc	aagatcaagt	acgaatgcaa	cgtttgcgcc	aagactttcg	11100
gccagctctc	caatctgaag	gtaggccttg	agagagagca	gtccaagggg	ctgtgagtgc	11160
atgcttgtgt	ttgtatttag	cttgctttcc	atgggggtat	gattgcattt	gcagtagtat	11220
gagccccggg	ttggggatag	tgggtatgga	ttccgcctgg	cttttgccac	ctctagctct	11280
ttgacttttg	acaagtgact	tcccttctcc	tgattttctt	ctgaataata	aaaaaattag	11340
gggtttggac	tagaagatta	ggtgaaactc	cctgctagcc	tgtgattttt	gtgcttttaa	11400
gaaaaacacc	attctgaaaa	catgaagatt	tcttcttttt	aagactgtct	tgatgctttt	11460
cttaagatat	ttgcatcaac	acttgagtct	tggagcagaa	atgttaggtc	tcagagccag	11520
cttgagagca	gagctaacac	atgtggcttc	ttcccaggtc	cacctgagag	tgcacagtgg	11580

agaacggcct	ttcaaatgtc	agacttgcaa	caagggcctt	actcagctcg	cccacctgca	11640
gaaacactac	ctggtacaca	cgggagaaaa	gccacatgaa	tgccaggtgc	gcagtatatt	11700
ctgggtagac	cttctgacct	ttgtagaaaa	tgtctgtgag	tcacctctcc	atgtctctata	11760
tagcccgtag	ttaaagccaa	caccagattc	tgcgttgtcc	catcctggac	tgatggcact	11820
atggtccttc	ccagtacttt	gtatctgctg	atgacttgag	atggcacagc	cagcttccag	11880
tgggtgggaa	aatggtaggg	gaaataaaca	gcccctcgtg	tgctgtgtgc	ccacatcccc	11940
ccgtttgctt	aataccacac	tggaggtgcc	acaaggaggc	ttctcacctc	ctaggttgct	12000
gggcgttggc	cggtaagcct	gcccctcccc	ttggcaactc	ttaatcttct	ggccttctctg	12060
tctcccttcc	ctgctgtctc	tctcccttac	actgtagggtc	tgccacaaga	gatttagcag	12120
caccagcaat	ctcaagacct	acctgcgact	ccattctgga	gagaaacccat	accaatgcaa	12180
ggtgtgccct	gccaaagtta	cccagtttgt	gcacctgaaa	ctgcacaagc	gtctgcacac	12240
ccgggagcgg	cccacaagt	gctcccagtg	ccacaagaac	tacatccatc	tctgtagcct	12300
caaggttcac	ctgaaaggga	actgcgctgc	ggccccggcg	cctgggctgc	ccttggaaga	12360
tctgacccga	atcaatgaag	aaatcgagaa	gtttgacatc	agtgacaatg	ctgaccggct	12420
cgaggacgtg	gaggatgaca	tcagtgtgat	ctctgtagtg	gagaaggaaa	ttctggccgt	12480
ggtcagaaaa	gagaaagaag	aaactggcct	gaaagtgtct	ttgcaaagaa	acatggggaa	12540
tggactcctc	tcctcaggtg	gcagccttta	tgagtcatca	gatctacccc	tcatagaagt	12600
gcctcccagc	aaccactac	ctctggtacc	tgtaaaagtc	aaacaagaaa	cagttgaacc	12660
aatggatcct	taagattttc	agaaaaact	tattttgttt	cttaagttat	gacttggtga	12720
gtcaggggtg	ctgtaggaag	tggcttgtac	ataatcccag	ctctgcaaag	ctctctcgac	12780
agcaaatgg	ttcccctcac	ctctggaatt	aaagaaggaa	ctccaaagtt	actgaaatct	12840
cagggcatga	acaaggcaaa	ggccatatat	atatatatat	atatatctgt	atacatatta	12900
tatatactta	tttacacctg	tgtctatata	tttgcccctg	tgtattttga	atattttgtg	12960
ggacatgttt	gcatagcctt	cccattacta	agactattac	ctagtataaa	ttattttttc	13020
aatgataatc	cttcataatt	tattatacaa	tttatcattc	agaaagcaat	aattaaaaaa	13080
gtttacaatg	actggaaaga	ttccttgtta	tttgagtata	aatgtatttt	tgtcttgtgg	13140
ccattctttg	tagataattt	ctgcacatct	gtataagtac	ctaagattta	gttaaacaaa	13200
tatatgactt	cagtcaacct	ctctctctaa	taatggtttg	aaaatgaggt	ttgggttaatt	13260
gocaaatgtt	gacagttgat	gtgttcattc	ctgggatcct	atcatttgaa	cagcattgta	13320
cataacttgg	gggtatgtgt	gcaggataac	ccaagaataa	cttaagtaga	agaaacaaga	13380
aagggaatct	tgtatatatt	tgttgatagt	tcatgttttt	ccccagcca	caattttacc	13440
ggaaggggtg	caggaaggct	ttaccaacct	gtctctccct	ccaaaagagc	agaatcctcc	13500
caccgccttg	ccctccccac	cgagtcctgt	ggccattcag	agcggccaca	tgacttttgc	13560
atccattgta	ttatcagaaa	atgtgaagaa	gaaaaaaatg	ccatgtttta	aaaccactgc	13620
gaaaatttcc	ccaaagcata	ggtggctttg	tgtgtgtgcg	atttgggggc	ttgagtcatt	13680
gtgggtgttt	gttgttgggt	tttgttgcct	tttttttttt	ttttttttta	atgtcaaaat	13740
tgcacaaaca	tgggtgctcta	ccaggaagga	ttcgaggtag	ataggctcag	gccacacttt	13800
aaaaacaaac	acacaaacaa	caaaaaacgg	gtattctagt	catcttgggg	taaaagcggg	13860
taatgaacat	tcctatcccc	aacacatcaa	ttgtattttt	tctgtaaaaa	tcagattttc	13920
ctcagtattt	gtgtttttac	attttatggt	taattttaatg	gaagatgaaa	gggcattgca	13980
aagttgttca	acaacagtta	cctcattgag	tgtgtccagt	agtgcaggaa	atgatgtctt	14040
atctaattgat	ttgcttctct	agaggagaaa	ccgagttaaat	gtgctccagc	aagatagact	14100
ttgtgttatt	ctatctttta	ttctgctaag	cccaaagatt	acatgtttgt	gttcaaagtg	14160
tagcaaaaaa	tgatgtatat	ttataaatct	atttatacca	ctatatcata	tgtatatata	14220
tttataacca	cttaaattgt	gagccaagcc	atgtaaaaga	tctacttttt	ctaagggcaa	14280
aaaaaaaaaa	aaaaaaaaaa	gaacactcct	ttctgagact	ttgcttaata	cttggtgacc	14340
tcacaatcac	gtcggtatga	ttgggcacct	ttgcttactg	taagagacct	taaaaccttg	14400
gtgcagtggt	ggggaccaca	aaacaaccag	ggaggaagag	atacatcatt	ttttagtatt	14460
aaggaccatc	taagacagct	ctattttttt	tttgccactt	tatgattatg	tggtcacacc	14520
caagtcacag	aaataaaaaa	ctgactttac	cgctgcgaatt	tttctgtttt	cctccctact	14580
aaatactgat	acattactcc	aatctatttt	ataattatat	ttgacatttt	gttcacatca	14640
actaatgttc	acctgtagaa	gagaacaaat	ttcgaataat	ccagggaaac	ccaagagcct	14700
tactgttctt	ctgtaaactt	caagactgac	agctttttat	gtatcagtg	ttgataaaca	14760
cagtccttaa	ctgaaggtaa	accaaagcat	cacgttgaca	ttagacaaaa	tactttttgat	14820
tcccaactac	tcgtttgttc	tttttctcct	tttgtgcttt	cccatagtga	gaatttttat	14880
aaagacttct	tgcttctctc	accatccatc	cttctctttt	ctgcctctta	catgtgaatg	14940
ttgagcccac	aatcaacagt	ggtttttatt	tttctctctc	tcaaagttaa	aactgaccaa	15000

<210> 66
 <211> 46340
 <212> ADN
 <213> Homo sapiens

<400> 66
 tatttttactt cagtaacaga aaatgaaaga aatgtttttaa tgttgctgat tgtattacct 60
 tcaggatcaa tagcagaagg acaaacttct ttgaggagat ctccctagtgt gtgcaactgt 120
 ccatctgcag ccacaggacg aaacagcttc tgaatgaaag gtcttttcagt cgttgtctat 180
 ttgaaaaagg aaaaaatgat tcaagcaatt aagtcctttgt tgctgccaat tacaaattta 240
 tatatcataa acttttatgtt ggcattaggt gcctttttgat acggtgttag cataattaca 300
 caacatcaca gatgtggtat cactgtgaaa aatgttttaac atgataaatt caggtaaatc 360
 taattctgag gaaacagaca aatccaaagt tgggtgggac attctaaaga taattggctg 420
 ggacccttca aaaactttaa gacattaaaa agcaaacaac acaaaaagat atcaacaaaa 480
 gcattttttc tcagtatctc ttaaagagac taacaaagca aatacaaaac ataaaccatg 540
 gctgaatact aaattgaaga aggacatttt ttagaaatcc aactatgaaa cacagttttg 600
 ggataaatgg ggaaatacag aatggacaac tgataatatt attgagttaa tgtcaaattt 660
 cttaggtaca ataaggacaa tctttatttt taagaaattc attgttcaag tgtttaggaa 720
 agaagtgcc tgaatccaa aacttaattc tctttctctt tttttggaga cagagtctcg 780
 ctctgccacc ccggctggag tgcagtggcg cgatctcagc tcaactgcaac ctctactttc 840
 caggttcaag tgattctcat ggctcagcct cccaagtagc tgggactaca ggagtgcgcc 900
 accatgtcca gctaactttt tgtattttta ctagagatgg ggtttcacca tgttgccagg 960
 gctgggtctca aactcctgag ctcaaggcaat ctgccggcct cggcctccca gagtgttagg 1020
 gttacaggcg tgagccaacc gctcctggcc ccaaaaactta accatctaatt ggttgagaga 1080
 gagacagaga gagagagaaa gagagagaca gagaatgtgt gtgtgtgtga agacaaagca 1140
 aaaataaaaa aatattaact aatgggtgatt ctaggtagag ggtgtatgat ttagtagatt 1200
 tcattatttc aacttttctga taggtttcac aattttccaaa acagcagatc cagccatttc 1260
 atctgacaaa aactgttagc agcactacat cgtaatttat tgctaataat ctcattgttt 1320
 tactcttaaa attgtttcat ttactaaatt tcttagtga tgatggaggc tttatcatga 1380
 cagagtacag aggctctgaa atgagccagt gtctatgaag agcaccactg tttgcaagat 1440
 ctatgatctt gtaccaggtt tcttttatct gttaatttgg gacattccat atctcttgag 1500
 tttgttgttg aaataaatga gcaactttgc caaccacaga gtaataaat aaatgttaaa 1560
 gagaataaaa gcatttttac ctctctctc cctcttaacg gttatttcac tttaaagatg 1620
 taaattttta gctttctgag atgaaaaatc attaaaaact aacaagaaca gagaaatgcc 1680
 atacatacat attttttggt tgcttggttc ctgagacaag gtttcaactc gtcaccagg 1740
 ttgaattgca gtggtgcaac cccaagttg caatcctcca cctaagcctc cagagtagct 1800
 gggactacag gtgtgagcca ccatgctcag ctaatttttt tacttttttg tagaaggggg 1860
 tctcactatg ttgccagggc tgcctcatat tttataagaa tatgacttca aacacttagg 1920
 cattagcgac aagggtttgt ttttgtcttt taatgacaga ggtatacctc aacatatttg 1980
 acacaactgt tagagatttg gtttaaaaag aaatagacat ggatgaagct ggaaactatc 2040
 attctcagca aactaacaca ggaacagaaa accaaacacc tcatgttctc actcacaact 2100
 gggagctgaa caacgagaac acatggacac aggcagggga acatcacaca ccaaggcctg 2160
 tcggggagta gggggctagg ggagggatag cattaggaga aatacctaac gtagatgagg 2220
 ggctgatggg tgcagcaaac caccatggca catgcatatc tatgtaacaa acctgcacat 2280
 tctgcacatg tattccagaa cttaaagtat aatacaaaa gaaaaataa ataaaaataa 2340
 gttagaaaaa taaacatgta agcatgtgag ctgcctttcc taattctatg tttatgtatt 2400
 cactgaatac atagtatttt aaaaatagtaa tccaataata tatttgagtg tttgtgacaa 2460
 gtatgaaaat tgtaattttt aaaaaatctt gataatatgc attgaatatg atttaattca 2520
 ctctactatt tgaactcttt agggattatt tttaaaaata tgattgatat cctttgatat 2580
 gttttggctc tgtgtttcca tccaaatctc atctcaaat gtaatcccca cccgtctagg 2640
 gagggactgt aatccccatg tgtcgaggga gggagtgatg tgggtcatag ggggtggttt 2700
 cctcatgttg ttctcgtgat actgagtgaa ttctcatgag atctgatggt tttaaaagt 2760
 gcagtttttc ctgcactctc atctctcttt cctgctggct tgtgaagggt cctgcttccc 2820
 tttctgccat gattttaagt ttctgaggc cccacaagc catacggaac tgtgagtcaa 2880
 ttaaaccctt tgccctttata aattatccag tctcagatat ttctttaaag cagagtga 2940
 acagactaat acattcttca atttaaaaag ccatactttc tcatacaagt tgaaaccaag 3000
 aacaatatca tgcataatca agtgattaac tgtgtaaaaga taataaggtt gaggagttca 3060

gagaagaaaa	gaaatgaata	gggaactgta	gtgataattt	aaaatagcca	tccctcactc	3120
agggtttttg	atcttcaggc	catgaagaag	cttttaatgc	tttttagcaa	aggaagtaac	3180
gttggtgaaa	ggctttttct	gacgactaat	ggaaagcagt	gctatgtatg	gtgacttggc	3240
tatgaaccaa	aaccagaatg	actggtgaga	ggctgactga	atacagcaag	cttatgtgaa	3300
gacaactgga	gctggtgcag	tggaaaagga	agacagcagg	actgtaccca	caactcaaa	3360
aaaaaagtca	gaaggtacct	cccgcagtc	aacctgaaaa	caacaaagtc	aaaggaatct	3420
tttcaagaat	ttggagctct	cattcatatc	ctaattagt	tatgaaatgt	gaggtggctt	3480
tgtataaatg	aaattacctg	gaatatttct	aacacaaaga	aataataaat	gcttgagggtg	3540
gtgaatatcc	tcatttgatc	attacacatt	gcatgcttat	agcaaaagat	tacatgtacc	3600
ccataaataa	ttgcaactat	tatgtatcca	taataattaa	aactaaaaga	ttaaaaatta	3660
cctgaaaaaa	aatgctaaac	aggaaaaggcc	aactagtctt	ggttacatac	taaaaaacag	3720
aaattcttct	ctaaccctac	tattggagaa	atatcctgtt	atttttatat	atcttttttt	3780
tcaccctttc	ccaaatctga	gcaagtatta	taaaaggata	accttcaaca	atcttttatg	3840
atgagggtatt	tgcttactgg	ggacaaaagcc	ccagtgcctat	tacatagtgt	agctaaacgc	3900
tgtagaatgg	taaaaaacaag	aaaatgctca	gcaaagtgtt	gtttctcatt	taatgaaaat	3960
cttatttttaa	aacacaaaaa	ctcaatatac	cccaaccaaa	aatctgatga	acattttctg	4020
tttaatatatt	attatacagt	acctttaaaa	acgtaatat	cttattctta	aaaatttagt	4080
gtgctagcaa	atagcaatta	agtacctaag	tcaatcagga	cgacaaaaaa	atactcaatt	4140
tggggagtta	gttacttcta	tcattctgaat	gcgtccctcc	aaaattcatg	ctgaaaccta	4200
ttcctcatca	tggcaggtatt	aagaggtgaa	gcctttgaga	ggtaattagg	tcatgagggc	4260
agagtcctca	agaatgggat	caatgctctt	ataaaagagg	ccccagggag	cttgtaaggc	4320
ttttgcccc	tctgccatgt	tgggggggtg	gggggtgggg	cgagcaacc	agtgtcaact	4380
ctgaagcaga	gagcagccct	caccagaaac	cgaatctgtt	gaagccttga	tctctgactt	4440
cccagcctcc	agaactgtga	gaaataattt	tctgttgttt	ataaattacc	cagtctaggc	4500
tgggcgtggg	ggatcacctg	aggtcaggag	ttcaagacca	gcctggccaa	tatggtgaaa	4560
ccccatctct	actaaaaata	cagaaaatta	gctgggcata	gttgtgggcg	cctgtaatcc	4620
cagctactca	ggaggctgag	gcaggagaat	cacttgaacc	cagaaggcag	aggttgcagt	4680
gaatcaagat	catgccattg	aactccagcc	tgggcaacaa	gagggaaact	gtctcaaaaa	4740
aaaaaaaaaa	aagtacacac	tctaacatat	tttggatatag	cagcccaaat	ggaatggact	4800
aagacaatta	cccttaaaat	aaaagctccc	atagagagat	catgcattca	agtaacagag	4860
ttcttaaggg	ggatgggaat	ggaggacata	ttcctgcaaa	cttttcaaca	gtctctatta	4920
gcccgatgtt	agagctctgc	aaagaagact	aaattatact	gagaaatatt	tttaaatctc	4980
cacaaatagg	aatgctgtaa	acgttgattt	agtatatata	aaattagaca	agactaacia	5040
tatccaatgc	aatctaaatc	ttagggtgac	agacaagaaa	gccactgcaa	acaggaatat	5100
accacaatac	ctgatcttgc	cacatatttg	taaataatgca	aagtatttca	ataacttcca	5160
agaaacagta	tactctcat	gagaaataag	atgatgtaag	tcacctttga	aactgtcctt	5220
gttacttttt	caaatgtatg	ttagtcattt	cttaacacca	aatgaaatga	aaaaactgag	5280
tggtaatggc	tggtgctctc	catctctcct	ctactcatgt	gccttcacca	atacagcaat	5340
cattttttct	tatatgggaa	atttacagtg	ttgatatagc	tcagagatat	attgaagaaa	5400
agcagaaaaa	cgaaacttat	aaacatttta	ggaaacctta	tgtattttct	taaatagttc	5460
aatgttaaaa	cttagaattc	ttataaataa	tgtgtgttac	agctatattg	taaatgggtg	5520
ctcatgcctg	taatcccagc	acttcaggag	accgaggtgg	gaggagagct	tgagcccatg	5580
agtttgagac	tcacccgggc	aacacagaga	gacctcatct	cttaaaaaaa	aaagaaagaa	5640
agaaagaaat	gaaatgcaaa	gaaaaagtct	ctattttcaa	tgtagccagt	agagccaata	5700
ggttaaccaa	tattaacatt	aacgttgata	aaacaagaaa	tgatgattta	ctataagctg	5760
aaaatcagac	aatgtatgga	ctttaagagt	aacaggcacg	atcatcacia	acttaaatca	5820
ggtttgagtc	ctatgagtta	tatacagtta	catgatgcaa	caaaagatgc	cagccagttg	5880
ttaaagagta	ttagattcgg	ctgggggtgc	tggctcatgc	ctgtaattcc	agcactttgg	5940
gaggccgagg	agggaggatc	acgaggtcgg	gagtcagaga	ccagcctggc	caatatagt	6000
aaacctgatc	tctactaaaa	atacaaaaac	tagtcaggca	tggtggcacg	tgctgttaat	6060
cccagctact	cgggaggctg	aggcaggaga	attgcttgaa	cccagggggc	ggaggttgca	6120
gtgagccgaa	atcgcgccac	tgcaactctag	cctgggcaac	agagcaagac	tctgtctcaa	6180
aaaagagtat	tagattcaag	tcctgtttct	gtcattttatt	atggaaccat	ggacacaact	6240
acctatcttt	cctgaacctc	agttttttca	actgcaaaac	aggaatatat	acatatgtgt	6300
atatatacat	ctgtgtaaac	acatatgtgt	atatatacat	ctgtgtaaac	acatatgtat	6360
atgtataaat	ggagataata	cctacattat	agtttctgag	ataataaaa	gcacaacaca	6420
attctgacac	ataacaattt	gtaacttaaa	acataccatc	accagggcca	ctagttttag	6480
aacactgtaa	tgcatagtct	aatttaatac	tatgcaaact	gtgttcactc	aaggttttat	6540

ttccttttaa	tttcattcat	ttactcttca	gttgtttgta	agctaaaaag	tccagaatca	6600
tgaaattcag	aagtttacgt	tttaatgttt	ttctatatgg	caaggaaaaa	aaaaagggca	6660
aagtcatttt	aacactactt	tcaaaatcag	cctagaactt	aacactaaaag	gcatgaccca	6720
taaaagggaa	tactaataaa	tagacttaat	taaaattaaa	caacaacaac	aacagctaaag	6780
cttttgttct	gcaaaaagatc	ctgtgaagag	aatgaaaaca	taagccgcag	gctgggagaa	6840
aatatttgca	aaccataattt	ccgagaaaag	tcttgtgtct	ataatatata	agaactccca	6900
aaattcaaca	gttttttaaaa	aaagcaataa	atccaattag	aaaatgggca	aaagacatga	6960
acagacattt	taccaaagag	aatatatagg	tggcaaataa	gcatatgaaa	acatatctca	7020
cacatcatta	gccattaaag	aatgcaaat	taaaaccaca	atgtgatata	attacacacc	7080
taccaaataa	tccaaaataa	aaattagtg	taacacccaa	tgctgggtgcg	catgtgggaa	7140
aatagtcctt	cacacactga	tggtaaaaa	gcaaaaacagt	acagtcctct	acgaaaggag	7200
tatggcagtt	tcttacaaaa	ctaaacatgc	acttaccata	tgaccaagta	attatactct	7260
tgaatatctc	cagaagtaaa	aatgtcttct	ccaaaaaact	tatacatgaa	cgttcatagc	7320
tgttttattc	gtgagagtca	aaaacagaaa	gcaatcccag	ggctacccat	taaaacaggt	7380
gaatgcttat	aaactgactg	taataggctc	gtcccacgga	atactactca	gcaataaaaa	7440
ggaacaaact	actgggtata	gcaacaactt	ggatagatct	caagggagtt	atgttatgtg	7500
aaaaaaagtca	atctcaaaa	gttacacact	gcatgactcc	actgatataa	cattagtgaa	7560
atgacaaaaa	ttttagaaat	ggaaaacaaa	ttagtagttg	tcagaggtta	gggaagaaat	7620
gcagtaaggt	aggtggctgt	ggctataaaa	gggtagccta	agagatcctt	ctgttgaaac	7680
gggtatatatt	tgaatatagg	gtgaatttac	atatgtgata	aagattgcat	agaactaaa-	7740
acacacacac	agtatatgta	aaactaagga	aatctgagta	aggtttgttg	attataattaa	7800
tacaatttcc	tggttgtgat	actgtactgt	aattatgcaa	gatgttagaa	ttgggggaaa	7860
ctagatgaag	ggtatgtaga	tctttctgta	ttatttctta	caattgcatg	tgaatctgta	7920
attatctcaa	aataaaaaatt	tttttcaaaa	tttcaaaaaca	actagtctag	agctttgtta	7980
atcaaagttt	tctctgagga	cctgtagcat	tttggttatc	acctggatct	tattaaaatg	8040
tagattctca	ggctgcatat	tggaaattcct	gaattggaat	ccgcatttta	acaagatttc	8100
caagtgattc	atgttttaaag	tttgagaagc	actagtctac	aacaatgact	tttaaccttt	8160
caacctactc	taacacactt	gaaggccata	acaaaattca	catcaataac	agttgctcgg	8220
ttggacagtg	actctcaaca	caaattgagt	aggaagggtg	gggactcaag	actcaggtag	8280
caggaaaagc	cccttaggtg	atcctgatga	aatgttttct	ccatcctggc	tgaaaaaccc	8340
agaacagtca	attaaggctc	aaaacaaaag	taatgtttat	aatactggag	atctttaaaa	8400
ggcagataat	atatactata	acagagcaaa	ggtaattatt	acaatgtata	aatctttata	8460
gaacccaaaat	cagaattaaa	atcactaagc	acataatgaa	aatcctttta	aaagtataaa	8520
aatgaattgta	gtctaagtaa	atactaataa	tggcagttat	agtgagaaaa	gctctagagt	8580
cttttactct	tcatacttcc	tagtcacaaa	catctatttc	caaaactgac	ccttcgtatt	8640
tcaaaataatt	tatggccttg	tacagtaata	agagcatgat	atttaaagcc	agtcagaaga	8700
cacatattct	agctctggat	ggcacttgat	gacgatggat	tcagcttatg	gttccaatcc	8760
cagctctgtc	aattagtagc	tatatgaccc	tagtcaaata	cttaaaccct	cttgtgttac	8820
tttgtgtgta	attgtatcat	ctataaaaat	aggatattaa	cagtatatat	ctcatagatt	8880
tttttgtgaa	ggttatacaa	ttaattcata	taaagtattt	agaacaatgt	ctagcacagt	8940
gaatttctcaa	tgagtgttat	aatgtgttct	tttaaatgtg	acttgactct	caacagaact	9000
ctactgaatt	ctaatatgta	ttctgtattg	agctgtcaaa	aaaaataagg	attataataa	9060
catatactat	tcttgtagtc	aacctgttta	ctatgttatt	actagtgtca	gttttgttgt	9120
tttggtcata	catattgttt	tacatacatt	agaattattt	agaaatgttg	gtttattaaa	9180
aatgaccatt	tatggctaga	agggtatata	tctggctcac	tgactgtgga	gtcaatgtcc	9240
ataaaagagga	ggaagaatgc	catcagagta	aaaggagatt	ctattcactg	aaacaaagt	9300
ataaaaagct	atgaaaagaga	aaaacataaa	ataaccaaa	gggtgaaact	taacagatgc	9360
ccagtagatg	cacaatgcac	tgggttgtaa	aacttaaaat	ggccttaatt	aaaagccaag	9420
cacggatgga	ggtgctggg	gagtctccta	cggacacagc	aggcagaatg	taacaatgac	9480
aaggggctca	agtttattta	aaaagagatt	ggacaggccg	ggcgtgggtg	ctcacgcctg	9540
taatcccagc	actttgggag	gctgaggcgg	gtggaatcat	aggtcgggag	ttcggaggca	9600
gcctggccaa	catggcgaaa	cctcatctct	cttaaaaaata	aaaaaaatta	gccgggagtg	9660
gtggcgtgca	tctgtagtcc	cagctactca	ggaggctgag	gcaggagaa	cacttgaacc	9720
tgggaggcaa	aggttgagc	gagctgagat	catgtcactg	cactccagcc	tgggcaacag	9780
agtgagactg	ctcaggatct	cccaaagacc	caaatccctg	taaactgaat	gcataatatc	9840
atttgctcca	gtgaggctta	gatggacatt	ctagtcttct	tgggtgagct	gaagaaacaa	9900
atattatatt	gataatttat	gtatgttgta	tttttcaagg	tatagcaaca	agtttttatt	9960
catcagctac	tttgtgtgtg	tgctttgttt	ttaagtcttt	tgaacacagga	tggtgattta	10020

ctacattttat	aagtaaaaatt	tattttgattt	acaaggggttg	cttaagtgtg	tcacaggatt	10080
tcacttggtta	tattttgcagg	tgcttaaaaa	atcagctata	ctaaactata	actggaatta	10140
gcaaagttca	tttattgatt	aatcaagaat	ataattagat	ttgcctaact	atataagtag	10200
tactatgtgt	tattttaagaa	ttaaacttag	aaaagggatg	gactctggaa	atatcaagaa	10260
gtgaaaaaga	ctgctctcat	ttttgtacaa	caattactaa	atcttctaagt	agcattaatt	10320
gaactgaaaa	ggcatttttag	aaaaactaga	ttttacaatt	tataactcta	ataaaacaca	10380
actaactatg	agtgtgcttg	ttcatgccca	aaagctacct	tccaaaaatta	aaaaccctat	10440
tggatggctg	ggtgcagagg	ctcatgcctg	taattccagc	actttgggag	gccaaaggcgg	10500
gcggatcacc	tgaggtcagg	agttcgagat	cagcctggcc	aatatggtga	acccgtctct	10560
aacaaaaata	caaaaattag	ccgggcgctg	tggcgggtgc	ttgtaatccc	agctactcgg	10620
gaggctgagg	caggagaatc	acttgatcct	gtgaggcgga	ggttgcaagt	agctgacacc	10680
gtcccactgc	actccagcct	ggcgagagc	ccagagcgag	actccgtata	ttaaaacaaa	10740
caaaacaaaa	ctcaaaaaac	cctattggca	attactaggg	ccatcaaadc	agtatatatt	10800
cacttgacac	acaattttga	gataatgaac	cgaacttact	atttttgaaa	atattacata	10860
ataaatatta	gtgaagcttc	attgctgaaa	tgggtgacaaa	gatgaatagc	aataaaactt	10920
ttcttataga	tcttttagcaa	aaacaaaaaa	accccaagca	tactatggta	cattacttta	10980
gagaatcaag	tagctgctag	ttgagtaata	gtggtaatag	gcactacaat	gatataaaca	11040
aattacaaca	aagaatattg	tttttatttc	ctgtccatgt	tttaaaaaag	ctttggtttt	11100
acctatgttt	aacaaaagca	taggtacaac	aacgactact	actactaaca	tataagtagc	11160
ctggatagaa	ttatcttaat	agtagtacct	aagtgcagga	tctctaagta	atgatcagaa	11220
ggcaggaata	aattttatca	gaaatcttca	tccattacat	atttactatg	catttaccag	11280
ggtatcacta	tgctaattgga	tacaaagata	aataacatgc	aaacaactgt	aatacagtgt	11340
tatgtgataa	cagaaatatg	tacaaagcac	tatgaaaaaa	attacaaaagc	ttgagcacaa	11400
attttaactc	tggaacttact	ggcattttaga	gcaaaaaccac	aacaatccta	actgggttaatt	11460
ttcattttct	aagagtttga	agctatatca	gtaggtacaa	agtaaaatat	gctaattgtg	11520
gtagaaagta	aaatattaca	acagtagaga	atttcaaaaag	aagataaaaa	taatggaggg	11580
aatatagaag	gtcttcaagc	ttccagcttg	aaatacatat	ttttttttta	atagagaaaag	11640
agataaagtc	atttgagtat	tcagagggca	gactgaatat	aatggtaact	ctgagaaatc	11700
agtggataag	gagagaaaag	tggactaaag	gccatagcat	atagagcttg	gaattgtcaa	11760
tgtatggtaa	ataacaaaag	tttgggttga	atcccaactc	ccaacaacgt	actgtgtatc	11820
tagagcaaat	tacatcaacc	tttgggagta	ctgtttctga	atctgaaaaa	tgaggaaaaac	11880
ttatctttga	acaattgatg	tgataattaa	atgagatata	tgaaatatct	aatgtaacaa	11940
gtgcttaaca	atgactagtt	cttttcatct	ctctcttgaa	ccattgtgaa	acgtagaacc	12000
aagaaaggta	acagtattta	gttggttacag	aacccattaa	gagagaataa	aaaataactg	12060
gtattctaac	ttcagtttcc	tttgaagtct	tgttaagtga	aataaatatt	atgtggcaca	12120
aagaaaaaga	aaacaggggt	ttacacagga	tatgctgccca	gactttacca	acaatgacac	12180
atgatatctg	cttcaactgt	cccatgcata	tttggcttraa	gatataattca	tgcatatcaa	12240
attttacatc	acatgggttt	caaaagaaga	ttcattaaaa	ttagcttaag	aatgtacaca	12300
atatacaata	cctcattaaa	taaaaagaac	agaccatttc	caaatgaatg	cttttagagc	12360
tttacagtaa	acagtctttt	ggtggtagaa	agaggggggaa	cagagagggg	agtgggtggg	12420
agtctgtagc	acttatcaga	ctacttttat	ccttttatgta	gagaaatagg	agagttgaaa	12480
ataagcactt	tctgtactta	tgttgagagt	ctgaagccca	cttttaatat	tcttgacaac	12540
actaaaaaat	aataattaac	atttgaaaag	ctgtcattat	tatagtcagg	gacacttaat	12600
ctccaaagga	gaagtttctt	aattgatact	atgattaaat	aaaagcatcc	atcagaatta	12660
tatccacaat	ctgggtttgga	gtttatgttt	tgtcttattt	aaattgttat	acttattata	12720
attctgtcta	gacagtgcga	aatgtacttt	gtcatacaaa	cacttgaggc	aaattttctt	12780
caaataagcg	caacactttg	tttctcttct	gtatcctttg	actgaataac	gtgtgggtaca	12840
gagaagtaat	acttcccttt	cttgggatcg	agatcaattt	gatgcttggt	ataagcccat	12900
ttacagaaca	aatggtattg	cttttaaaatt	tttatatgaa	cttatcagta	gactagccaa	12960
aaaagaagct	tcatataaaa	gtgctaggat	tgatattctt	agtaataatt	aggtaaaattc	13020
tctaaaaatt	tctcccaaaa	gatctgaaaa	atcataccaa	gggaagtata	gttttaaaatt	13080
cattatataat	aatagcttta	aaatatcttt	gctaattcta	cccaaagcca	cactaaaaaag	13140
actaatacaa	aaagaatgta	attaataaac	tttttctctc	tgaagaatca	aggggcactt	13200
ctgcatatga	acatgtttta	tctttttggt	gtactttacat	aaaataatta	agaaacactt	13260
ttaattagta	taaacaaaga	aatcaaaaata	gcaagaagaa	atgtctgagt	aaaagcagct	13320
gtgctgacct	caaaagtgaa	attctgttct	cttgatgccc	agttaagtgt	ctaaccagag	13380
gaaaagtgat	tctaaacctg	ggctaggagc	tagtggagct	cttcaaacag	tctcacctac	13440
cctcaccctt	caaggaatgg	tctatgggtt	ctgtggtgaa	cgctaaagtt	tataacatgg	13500

gaatattttat	tatttttgttt	ctaacacaaaa	taattttttaa	aaattttattc	tactaaaagta	13560
acatcaaagg	gaaattttcat	aaaaatttctt	ttgaaattttt	tagaagtagc	aaataaaggc	13620
aagtgataaa	tatttttacag	atttcaccac	ttacgtaatc	tgatcaacaa	attttaaaaa	13680
catagcactt	gaatactatt	aaaaatatat	taaaaaggt	acatagtaaa	actataaaaa	13740
tcttttaaaaa	aaatataaga	ggaaaccttc	gtgaccttgg	attaggaat	ggtttcttca	13800
atcgggcaac	ctaaaaatac	aagcaaccaa	agaaaaaaac	agacaaaactg	gacttcatca	13860
aagttaaaaa	cttttgttct	tcaaatgaca	tcatcaagaa	aataaatccc	acagaatggg	13920
acaaaatatt	tgcaaacccat	atctgataag	agaccactat	tcagaatatg	taaagaattt	13980
gtaaaactta	taaaataaaaa	gttaaagaag	tcaatttttaa	aatgagcaaa	ggatctgaa	14040
acaatttctc	taagaaatac	gaatggctag	ttaaatgcat	gaaaagatgt	ttagcatcac	14100
tggtcattag	gaaagagcaa	aaacccaaaat	gatatactcc	ttcataccca	ctaagactgc	14160
tgtaattaaa	actatagaaa	ataagcgttg	gcaaggatgt	ggacaaaattg	gaaccctctc	14220
catacactga	tggtagaaat	gtaaaatggt	gcagatgctt	tggaacacag	tctgacaata	14280
cccccagggt	ttaaactgtg	aattaccatg	caacccagca	attctactcc	taagtatcta	14340
cccaagagaa	atgaaaatat	atgttcacca	aaacatttgt	acataaaatat	taactgcagc	14400
ttttattcat	aatagccaaa	aagtggagac	aatccacatg	tctatcaatt	ggtgaattga	14460
taaacaaaaa	gtggtatctt	catacaacta	ttactgggcc	ataaaaagaa	tgatgtattg	14520
atacatgcta	caaaaatgaat	gaaccttaaa	aacaatatgc	aagcaaaaaga	agccagacac	14580
aaaaggccat	atattacatg	atgctaatta	cataaaatgt	ccagaaggga	gaaataaaat	14640
agtagttgcc	aagggtctga	gggaggggga	atgatataag	tgactgccaa	tggtcatggg	14700
gtttcttttt	aggggtgatga	aaatgttctg	aaattttatc	acgggaatgg	ttgcacaact	14760
ctgtgtaact	tagaattcag	tgactcctaa	aaccaatgaa	tagcatgctt	taaaaggtga	14820
cctttgtctg	gcatagtggc	tatagtccca	gctacttggg	aagctgaggg	aagaggatca	14880
cttgagccag	gagttccagg	ctgtactgca	ctatgatcat	acctgtaaat	agccaccata	14940
cacaccagcc	tggtgcaacac	agaccatgtc	tctaaataaa	taaacaaaata	aataaaataaa	15000
aggggtgacct	ctgtagtatt	gagattatac	ttcaagtaag	ctgttattaa	aaaaaaaaaa	15060
gttatcatat	gggtggcagg	ggaaatcatt	ctgggatgat	ggctaacttc	atcagtattt	15120
gatttatacc	tatgcatcat	accttatgtt	tgttttatgc	attttgtggg	ttttttaaaa	15180
aaattatatt	tcataaaaaa	aaatttttaa	aaaattaaag	tcaagaacct	caaaacacaa	15240
aagatcagag	agacatttct	accttatcaa	ttcagaaaaa	ttacaagtct	ttttcttaaa	15300
aattgtatag	catcatggtg	attttaagtt	acctgtagga	atttaataaa	ctttgtctta	15360
actgttcacc	aaaactcatt	taatatccat	gttctgatac	tgaaaatgaa	gctgaaaagt	15420
tttgaaatta	caatatgcta	gtttaaaaag	gtttactaaa	atacataatt	tcattataag	15480
gagtaatatg	aaataaaaagt	atcaaatatg	ggaccattaa	aaatgtcctt	actaacaaa	15540
tgctaccac	attgtggact	cactgcgtcc	actgttttgc	agcttttcca	gaacgctcgc	15600
caccagttag	ggtagccaag	aactcctcat	cttcactttc	ttcctcacta	gcttgggaac	15660
tctggattcc	cacccacact	gctgtgacct	gaatggggaa	gagaaacgcc	atagtaaggg	15720
aactcttcct	tttatagatt	tctgaattag	aatctggcat	tacaaaagaa	caatgttata	15780
aatccagggt	agagtttata	gttctatttc	actattactt	atatggcttg	tcctaggaac	15840
ttaactatta	tttacaatgt	aagtacctat	ttccacaaaa	aaattcaaaa	ttttggaata	15900
caatatctga	agagagaatg	gtctattgaa	ttccaaagtag	gctgatacat	cccaacagta	15960
tttcagattg	agataataat	aataccacca	attcatcaag	tcaaattata	tgcttatttt	16020
ccacaatgga	agtttttaaa	tagtataaac	attttaatat	atagcaggct	taacttatga	16080
ttattaaaca	gggttctaag	aaaatagtat	acatcaaaata	ttaatgtgct	tcttgtataa	16140
tttaggtgac	aatttatcca	tctgagaaat	gcaaaagaga	ctttggtaag	gggttgagta	16200
aggagcattc	tgtgtcaaag	aattcactag	caaaagaggg	tatactgtag	ttacaagcta	16260
taatcactgt	acttattttta	aatccctctt	cagaaccagg	tcttaaaaaga	tgataaacat	16320
ggcctcatga	ataactatca	accaaactat	agaaaagagt	gcaagagtgt	ggtgttctaa	16380
cttaaaatat	ggtgtttttat	tcaataaatt	ttattttaagg	ctccaaaagc	agcagcctca	16440
ttccccagaa	atcatagtta	aatgaaatct	tccttactaa	aggaaaaatg	aatcacata	16500
tttaacgtga	acatttttaa	aacactctaa	agcaacaaaa	ctattcaatt	gtatgtgata	16560
tggttagtaa	aggcatgtag	gtaaaaagga	ctaaaaactc	taataatggt	tgggccaaaa	16620
gtaaatttgt	tagttctact	ccattaagca	ttcctcaagc	agtgtaaaaa	ccagagttca	16680
agttacactt	tgatgtgtag	atcctttgaa	agccactcta	ccctgtttta	tatgaagcat	16740
ccgcagctaa	aatgaacacc	tagtgaagag	tatgaatgct	gcaatacata	agcagacgtc	16800
agaattgtcc	caagctgatt	ctaagttact	ttaaactatgt	atgcagagtc	agaatatgac	16860
ttacttctta	gaagtaacag	ataattacct	ttggcataat	gaaaaaaact	ttaaatgtaa	16920
gttaatacac	gtattttccc	tttagcaaa	ctttgctttt	aaaagaaaa	ttcaaaactt	16980

aaattaaaat	aggaaatgct	ctactatgta	gtaaaaatac	tttttagatt	actgaagcaa	17040
agaaaaggaa	ggattctatg	aggaggagaaa	agtgggagaa	aaatgtaaaag	aaaaaaaggaa	17100
agaaggaaaag	aaaagagaaa	aggaggaaaag	aacacaagga	cagaaaaggcc	tattgaaata	17160
tattattttct	ttcaaatttt	aaacgagcag	aataaattct	tttgttttat	aactatgaaa	17220
taatctatgt	tcctcttatc	tatgcttgga	aaatttagac	aaaatgttaa	gagtaagtac	17280
tacattggat	ttccgggtct	tcagctctga	aaacaagctg	tttcttaaca	tacgtcaatt	17340
ttctatattt	catgtcattt	ctatttgcaa	atgtttataa	gttcaatatg	atgtaaaaaca	17400
tggttaaatg	aagttcaaaa	ataagtataa	catacattag	tttggctatt	ccaaatttca	17460
tgcacattaa	ctcagccaca	catctaacac	agtcagccct	ccctatccag	gggttctgca	17520
tctgcagatt	caactaacca	tgggtcgaaa	atgtttttgt	accaaaccatg	tacaggcttt	17580
ttttcttggt	atcatccct	aactacagta	taacaactat	tttcacagtg	tgtacatgtg	17640
tatgaaatat	tataagtaat	ctacagataa	tttaaagtat	acaagagggt	atgcatagggt	17700
tatatgcaaa	tactacacca	ttttatatca	gactctcaaa	catcagtaga	atttggtaac	17760
ccaggggagg	cctggaacta	atcacccaga	ggtatcgaca	gatggctata	tataaatcac	17820
tcagtgaatt	caggattcac	attatttcac	aactagtata	attttatggt	gttcacataa	17880
ttgtgtcaca	acatatacat	gcagacaggt	gactttcatg	aaaagattac	acccaagata	17940
gacatatggt	ctactcaaat	acggtttcca	aatgtgtatc	caatcttggt	taattataat	18000
caaactcacc	attccattga	taagcgacct	ctaccaacct	gcttatcccc	tccaagcaat	18060
ataacagtg	ttctctgaac	caatattgac	cctcctttaa	attgatagcc	tttttttaaa	18120
aagctaacca	ttgagaagta	catactgttg	aagacagAAC	atattctgta	aaatgctccc	18180
aagatatcaa	agtcagatga	tacaactgaa	tgtttatgct	agattatatt	tctaagctga	18240
gaattacatt	ttaatatacc	ataagcaatc	tgcaaaagaa	gcaacttgcc	taaagatttc	18300
aggagtttca	agtatgcata	tgtcaatatc	tgtatcaata	tgtaatatca	atataatcaa	18360
tgcacacaac	aatacgtaac	tgtacttata	tcactctcct	agcactaatt	attacaaaca	18420
atctgcattg	actgcaaaagc	aaaagtataa	tataaaatcc	caaaaaacct	tgaaaattta	18480
ataaaaccaa	aaaacaggca	tcacacacaa	gaactgaggc	gtatacttca	ttaatgagta	18540
tgatatcctg	atatgaaatg	tcaaacaaaa	ttaccagggc	tcagggttaga	aataaagata	18600
ggacattagt	ctttgtattt	ttaaattgat	ttttcttct	aatattcctt	aatgataacc	18660
ctatatatta	cctacttaaa	attatttagca	aatagttatt	ttaaaagtat	gagtaattag	18720
accaaagca	ctctcatat	ttaccacaaa	gaaggaaact	ctaccaagaa	ctaaagccta	18780
gtaattctgt	tcttaacaga	cagggtgttg	gtattctggc	atgttacatg	aaaatcactt	18840
atgagaagaa	cagaaaaaaa	aattagaagg	tagttttcac	tatggaaata	ggtaagtgat	18900
taagcagatt	ttcttacacc	atgaaattgt	cagcagactc	aataatcacc	ctaaggggca	18960
tcattctgga	tgccgacatt	ctctatgatg	gaaagggact	gaaagtataa	tgcactaatg	19020
acataaagaa	accaatatcc	aatagtaaag	ttgaagaaat	aaacattctt	tggacaggaa	19080
ctaagctgaa	gttttgcaact	accaagaaatg	tattatgcc	gcagtaaat	aggaactaa	19140
agcccatgtc	aaccaatgaa	aaatgggagg	actgaaatca	atcattaaag	cagcagcaag	19200
gttctaacta	ttctaaggta	taggctacct	ctggcgata	ttatcagagt	tgacaattct	19260
tccaagaaat	tctaacatca	actgtaatct	gaggctcctt	aaaaataat	ataaaccagg	19320
cagtagactt	acatttttga	atattttctt	ctaagagctg	tacattaaga	ttttatttgt	19380
gatataaata	ctatcaaata	attagctata	gaacagctct	attttcaaca	gttataacat	19440
tttaagccat	ctcacattta	acctaaactt	ttatcaaagt	tcaaaactga	ggccgggtac	19500
ggtggctaac	acctgtagtc	ccagcaactt	gggaggccaa	gatgggcgga	tcacttgagc	19560
ccagggaattc	gagaccaacc	tgggcaacat	ggtgaaaccc	catctctata	aaaaatacaa	19620
aaattagctg	cgctgggtg	tgtgcgcctg	tagtcccagc	tactagagag	gctgagggag	19680
gagaatcacc	agggcctggg	agatcaaagc	tgcagtggag	tgagatcggt	ccactgcact	19740
ccaccctggg	tgacagagtg	agaccctgtc	tcaaaaaaaa	aaaaaaaaag	aaagaaagaa	19800
aaaaaatca	aaactgatca	cttgaggctc	aacttatggt	tactatatct	acttatattc	19860
ccaaagacat	cttaaggaga	gatgaaatca	taaaaagggt	aggatgagaa	agaaaatagt	19920
aagtcagtaa	ggtcaatttt	tacatatatt	aggctagcat	aataaaaaata	tgagtgtctt	19980
attattattt	ttttttgaga	cagagtcttg	ctctgttgcc	caggctggag	tgcagtgggt	20040
caatcatggc	ttactgcaat	gtctgccttc	caggttcaag	caatccttgt	gcctcagcct	20100
cctgagtagc	tgggattaca	ggtgtgcgtc	accctggccc	gctaattttt	gtatttttcag	20160
tagagacagg	gtttcaccac	gttaaaccat	cagtttggcc	aggatgggtc	caaactccca	20220
aagtgttagg	attacatgcy	tgagccactg	gctctggcct	aaagtgtctt	attataacca	20280
agaatttatt	tgtggagaga	ggtaaagaaa	actcattttt	agtgaataaa	ttaaaactgc	20340
atcattcaca	atctatcttt	caaaatgagg	tattaactat	tttggcttct	aaaattaccc	20400
catatactac	atgcatgagc	atgggaattg	aagttatttt	attcctaagt	ttgagacttc	20460

atgttttaat	gtgatcacta	aaaatttctc	aattgatgat	taggaaaata	actttctgta	20520
aaattccaga	attttagctg	tttcaatctc	ttcatattaa	ggggagaaca	ttatgttttt	20580
actttctgtg	catgcacttt	ctttattaga	agaaaatgga	ctgagggcag	taagcaaccg	20640
aaaaggaaga	gtaataagaa	gcctgatgtg	tgtgaaaact	ggagaacagt	ctcaaactcat	20700
aaaaagttat	gacagaagag	gcataaaaaa	taaaagtaat	gaacttaata	tatgaaagggt	20760
aataatgatt	aagagcatag	gctataaaagc	cagactggac	tccctggatt	caaactcctgg	20820
ctcttcta	tactaggtag	gtaaccctga	gcaagtttca	atgaccaatc	tttttctcaa	20880
ttacctcagg	tatataaaagg	ggacagtaac	agcattttaac	ccagaggaca	ataaggatta	20940
aataaataca	tgtaaaaata	tttaaaacag	tacctggtat	tcaataaagc	gcaataaatg	21000
ttagctgcta	ttattattca	tctaaacttt	actttcatta	ccagcaatat	tttttaactct	21060
taaaaatatt	gaataaaaaca	atgacctagc	ttagtaaata	aattcataat	gagaaaatgt	21120
tgatttcatt	taataataac	tttagtagtt	tgggataaca	ctttgcata	tttaatttcc	21180
ccagctcata	ataactcaaa	taatttgcca	tcagatgata	tgttattttg	aagttaacaa	21240
ataaagcatt	tcctaaaaaa	gttctaatac	ataacttttg	ctctcatctt	atgttttaaa	21300
aacaaaatgg	caaactcatct	gcatcaaaata	gttcttactc	ttataacatg	acaattgttt	21360
taaaaatata	ctgctggaaa	aagcaactga	agtcctagaa	aatagaaatg	taatttttaa	21420
ctattccaat	aaagctggag	gaggaagggg	aaaaacatat	ctgccaaata	agcttataat	21480
taatagtgtg	tttcagtttt	caaaaatcca	cataggaagc	aatttaagcc	taaaattgct	21540
aagtcata	ctcagcgtag	tagatagctt	agggcaatca	aaacttgctg	tgttgggctg	21600
ccccctacag	gactcaattt	acctatttct	tttaaaagggt	gtgtaagtag	gaaatgat	21660
tcaagtttta	cattaacaat	attaatgcta	aagcagatga	ttatcattca	cgcattcact	21720
ataggaggaa	acagtctctg	agaaccatct	atagagatac	agagagaaat	gaaacaatcc	21780
ttgtccttga	ggaatttaata	gtttactgct	tacagagaaa	ctacatacat	ggtgaaatat	21840
ttaaaaatag	ctcatgatat	cctctatgat	attatgtttg	ctatagaaaa	agaacaaggc	21900
tgaagatcta	agatccaagt	tctactgttg	gctctgccat	caaacaataa	gctaaacaat	21960
gtacaagtca	gttttgggga	agctgtctta	ttcccaaaat	gaggagggtta	aatttagtta	22020
ttcttcacag	ctctatggct	ctaataattcc	acagttacat	ttgtcaaaac	aaaaggtaga	22080
aggaaatggt	tcaaaaacag	acttcgcaga	aagaacatct	atatgatatg	aagggtctgg	22140
gcatatgtga	agaaatcaag	gaagacttct	tgaggaaggt	gacatctgaa	gtaactttag	22200
aagcactctg	ggagccaagg	ctattcccag	gagtttaacag	agtcagataa	taaaagatca	22260
aagatgttta	ggggaatagc	atgcagtggt	atttggttgc	agtctagcta	tatttttagga	22320
aacatcaaat	taatatcagt	ataaaactca	acagaatgga	gggagaaaaa	gcaggtagaa	22380
aaatctaaga	accactaaaa	tagttcatct	agaagataaa	ggacccatga	gctaaatcag	22440
tgcaaatggc	aagaagggaa	taaatgaaga	cagttctggt	ccattagaac	tgcaactcaa	22500
caaaagtgat	caaaagaggt	attccaaagt	attgacctgg	taacttgaag	aaaagtaaa	22560
aaagaggaaa	ctggacactg	aaacagaaga	agtagattat	gtatttggtg	gtgaattgaa	22620
gtagattggt	gggaccagtt	agaacctcac	agagaagaac	tatgttaaga	ccagaaatag	22680
ggccaggtgc	ggtggctcat	gcctgtaatc	ccagcacttt	gggaggcctg	ggtgggcgga	22740
tcacctgagg	tcaggagttc	aagaccagcc	tgacaaagat	ggagaaaacc	tgtctcccct	22800
gtctgtacta	atacaaaatt	agccaggtgt	ggtggtgcat	gcctgtaatc	ccagctactc	22860
aggaggctga	ggtaggagaa	tcgcttgaac	ccgggaggcg	gaggttgacg	tgagctgaga	22920
tcgcaccatt	gcactccagg	ctgggcaaaa	agagcgaaac	tcttgtctca	aaaaacaaac	22980
aaacaaaaca	aaacaaaaca	cagaaataca	tcaattaaaa	aagttagcta	ttcaccagat	23040
atgttccact	ggtcataaaa	caaaaagaata	caggaggcat	gacaagccat	catcattgct	23100
gttaaaaata	ctcacagcaa	aattataatg	atttaagtca	ataacatcta	ataattccag	23160
ctatagtgtg	caattttaatt	tattatgtgc	caggcacat	agtttattaa	aggtattacc	23220
tctaattttc	acaataacc	tattttacag	attataaaat	ggaggcccag	agatgtaagg	23280
tgaacgagcc	aatcaccta	gttacctgga	atataaactc	agaactgcct	aatcaaaaag	23340
ctctcaatct	taaccacatg	ctatactgat	gcatgtcaaa	gattcaattc	attcagattt	23400
ttcaagggtta	tcggaaaacc	tatgtagata	aaaattttcca	aaataatcaa	ggatatgtaa	23460
cttttacaga	aagcaatcac	tgatcatcta	ttgcaatact	catgttctta	agcaatatac	23520
tgagttgaaa	tttttatatt	ttataaataa	ttagaaagaa	tacatttttt	aaaactttta	23580
aaaacacctc	agtttttatt	ctcttcccc	aatttcaaca	aaatccattt	atccaaactt	23640
gaggttgatg	cattaaagtg	gtgatatcat	cagtaaatagc	agagttagga	ccctgaatat	23700
actctcctcc	ataaaaagcaa	caagaacaca	aaaatttctca	aaatgaactt	tttctgaat	23760
ctttcaaaaag	ccccactctc	agaaaactgt	cattattttga	tctgccagtt	ccctagaaaa	23820
acctccctca	taggacatta	tttgacttga	ctcagagctc	actcagtgca	aacaatttta	23880
tcaccaggag	agtttgtgga	aaatcagtg	caattgttaa	acatcacatc	tgccatgaga	23940

tagcaataac	agatgggaca	aacaagctaa	ccaaaaaatt	aaaagaaaaa	cctgggaaa-	24000
aagaaatcca	aaggggggtct	gaaaagtctt	aacatatctt	tgataatcca	gaaagccata	24060
cacatgtata	gagctgtgta	cacgctcaaa	aaacatctac	gaaggcccta	aactctcacc	24120
tatgggaaac	cctgaggctc	tgtacaagaa	gaaagtaaaa	tccagttata	aattgcttgc	24180
cgtatcattg	aaggcaatgc	cccaacattc	acacataggg	ccctggcaca	gattggaaga	24240
tactctagtt	ctaggcattc	aagaaaaatct	cttctaatac	tcagatgata	actaaactca	24300
ccaagcagta	acttttagggg	cctgtgtgat	aaaaaataaa	aacctgaaag	aatttagttca	24360
ggaaagaaac	taaacaagca	acagcaacaa	caaaaacaga	ccttgggaaa	ggggggaagc	24420
atctggtttc	cagagttatt	ctgttatata	atataaaaata	ttcaggtctc	aacaacaaca	24480
aaattacaaa	gacatgcaaa	gaaacaagta	taagccacaa	actgggggga	aaaagcagca	24540
gaaactggcc	ctgaaaaaga	ccagatgctg	gacttactgg	acaaagactt	taagagagtt	24600
atcttaataa	tgcgcaaaaga	actaaaaaaa	agtttatcta	aagaactaca	ggaaagtata	24660
agaacaatat	ttctgatcct	tcagaagaac	cactttttgt	cactacagat	tagttctgtc	24720
tggtctagaa	cttcttaaaa	acagaatcat	agagtataat	ctctttatat	cagctctttt	24780
tactcaacac	aatgttgtgt	gagatttatc	catgttgttg	catgtatcat	tcccaaacag	24840
aaatagaaat	tatagagata	aataggagtt	acaaaaaagt	accaaacaaa	aattctggag	24900
ttgaaaagca	caaaaactga	attaacttga	ggggtcaca	agctgatttg	ggcagccaga	24960
agaatgaatc	agcaaatcta	aagataggtc	aattgctgga	aagagaggga	agaaggaaag	25020
aaggaaggaa	aggaggctca	gagacccaag	agacaccatc	aggcatacca	atatacatat	25080
aatgagaggc	ccagaagaag	atgcagaaaa	agggtcagag	tatctgaaaa	aataatggcc	25140
ctaaacttcc	cgaacttgac	cccaaaaaatt	aatctacaca	tccaagaaga	taaacaaaac	25200
aaaaagaata	aaatcaaagc	gatccacacc	taggtacatc	ataatcaaat	gactgaaata	25260
taaagagaga	ctctcaaaac	aggcaaggga	cttatgtaca	aaacatcttc	agattaataa	25320
caaatttctc	atcagaaatg	atgttgtcaa	taggcaatca	gatgacataa	tcaaagcact	25380
gaaagaagta	gaatgtctgg	gacctggaat	gctgggtggac	acctgtaatc	tcagtatttt	25440
gggtggccaa	ggtgggagga	tcacttgagg	caaggagtgt	aagaccagcc	tgggcagcag	25500
aaagaggctc	tgtctctaca	aagaataaaa	agattggctg	aatgtgggtg	tgtggacctg	25560
tagtcccagc	tactcaggcg	gctaagggtg	aaagatcgct	tgagcccagg	agttggaggc	25620
tgcaactgtg	tatgactgtg	ccactgcact	cttgcatgtg	agaccctgtc	tctataagga	25680
aaaaatgtca	accaaactga	acatgcagaa	aaactgcact	tcaagaaatg	atcagtacct	25740
tgaagctctg	aagggtgctta	agactgtaga	tcaataccat	agaaaataat	ttagtattta	25800
ggaatgtaag	aaaattaaga	cagccttggt	tgataactac	acataatact	gtaactgttc	25860
ttgcaactgt	ctgggtattg	tcaagctatg	agcacaaaact	gatgactgaa	atacagaata	25920
cagaacagga	tataaaatct	tatcaggtaa	agttaggcaa	gcaattacta	gttgtaattc	25980
aacttgaaag	agaaggaaata	aggaaccaac	tcaaacagg	cagcaatgaa	ttgtaaaaaa	26040
gcttaaggta	aaacaaacag	ggaaataaaa	caactcagaa	cctaagcata	tcgtaagaac	26100
ctaacttaac	aaggaggggc	ttaaactgat	tatttttacag	cttgggtgca	attatcccac	26160
aaaaaacttt	caggagtttc	accagtcct	aaactatttg	gttattagaa	aatagcttta	26220
ttgggctacc	ctctttgggt	cccctccctt	tgtatgggag	ctctgttttc	actctattaa	26280
atcttgcaac	tgcactcttc	tggtccgtgt	tgtttacggc	tcgagctgag	ctttcactct	26340
ccatccacca	ctgctgtttg	ccgccatcgc	aggcctgcca	ctgacttcca	tccctctgga	26400
tctagcaggg	tgtccgttgt	gctcctgata	cagtgagacg	cccattgccg	atcccgaactg	26460
ggctaaagac	ttgccattgt	tctacgcgg	ctaagtcccc	gggttcaccc	taattgagct	26520
gaacactagt	cactgggttc	cacggttctc	ttctgtgacc	cgtggcttct	aatagagcta	26580
taacactcac	cgcgtggccc	aagattccat	ttattggaat	ccatgaggcc	agaaccccca	26640
ggtcagagaa	cacgaggctt	gccatcatct	tagaagcagc	ccgccaccat	cttcggagtt	26700
ctgggagcaa	ggaccccctg	gtaacaattt	ggcgaccaca	aagggaacctg	aaccgcgaac	26760
catgaaggga	tctccaaagc	ggtaatatgt	gaccactttt	gcttgctact	ctggcctatc	26820
ccttagaatt	ggaggaaaat	actgggcacc	tgtcggccgg	ttaaaaaacga	ttagcatggc	26880
cgccagactt	tagactcagg	tatgaggcta	tctggggaag	ggctttctaa	caaccctcaa	26940
cccttctggg	ttgggaacct	tggtctgcct	ggagccagct	tccactttca	attttcctgg	27000
ggaagccaag	ggctgactag	aggcagaaa	ctgtcgtccc	gaactcccgg	cattagccgg	27060
ttgagatcat	gtcgcagcca	gaagtctcta	ctcaacagtc	gccatgcgt	gcgtcctac	27120
cttcccttct	gtccacacac	tctgggtcc	caaccacgac	tttcttgaaa	gtgtagcccc	27180
aaaattctcc	ttacctctga	atctacttcc	tctgatccct	gcctcctagg	tactaatggg	27240
tgagactttc	atttccctcta	gcaagttgta	tctccaaagg	gatctaagga	agctctatgc	27300
tgcgccttta	ggcatctagg	ctataaaccc	agggagctct	gtccctgggtg	tccctcctga	27360
tttaggtata	cagctctaga	catgggcagt	tatgtgggac	ctgttccccca	ccacccttgc	27420

cagggcccca	agtttgtaaa	tggctaagag	aggaaacaga	gagagacaga	gagaaagaga	27480
cagtggagaga	cagacagaga	cagagagaga	gagagacaga	gaggagagag	agagagacag	27540
ggaggacagg	gagagagaca	gagaggagag	ggagagagac	aaagaggaga	aagaggcaga	27600
gagacaaaca	gggagtcaga	gaaagaaaaga	caaagataga	aatagtaaaa	aaaaacagt	27660
tgccctattc	ctttaaaagc	cagggtaaat	gtaaaacctt	taattgataa	ttgaaggtct	27720
tctccgcgac	ctataaacac	tccaatacta	ccttgtgtgc	agcgtaaaca	agggcgtagc	27780
ctgaaaaacac	taagaccact	gacaacccat	agccttccta	tcaaaaatcc	ttaacatcca	27840
gtgacctgcg	gatggcccaa	atgcattcaa	tctgtagcgg	caactgcttt	gctaacagaa	27900
aaaagtagaa	aagtaacttt	tagaggaaac	ctcattgtga	gcacacctca	ccggttcaga	27960
attatttctaa	gtcaaaaaag	caaaaaaggt	gcttattaac	tcaaaaatat	taaagtatgg	28020
ggctattctg	tcagaaaaag	gtaatttaac	actaaccact	gataattccc	ttaaccctgc	28080
agatttcctt	acaggggatt	taaatcttaa	ttaccataca	aaggtccgac	cagacctagg	28140
aggaaactccc	gatgatagat	gttctctccc	aaatgactga	ggaaaaaacc	ggaaaaaacc	28200
acaatgggta	ttcagtaatt	gatagggaga	ctcttgtgga	agcagagtta	gaaaaattgc	28260
ctaataattg	gtctctcaa	atgtcagagc	tgtttgcact	cagccaagcc	ttaacgtact	28320
taccgaatca	aaaagactat	ctcaatcctg	actcaaaagc	ttacttatac	cctctctgaa	28380
acgaattttgc	ctaagaactg	ttgtttatgg	gaatgcactt	tgatggagca	gctgggttgt	28440
tatgaaatac	tcaggaactc	agcctagctc	taggactcac	ccctgagcac	aaaggcaatg	28500
ttgggacacgc	tggtaaagga	ccactagaat	ccagcagccc	ggaccccttt	ctttgtgatc	28560
aagaaaaggcg	ggaaaaaggg	tgagggtctgc	tacatcagtg	agcataacta	atccgataag	28620
cagaggtcca	tgggtgggta	cacaccccg	aaaggaaata	gcattaggac	catagaggac	28680
gctctaggac	taatgctcat	cggaaaatga	ctagtgggtgc	tggtatccct	atgttctttt	28740
ttcagatagg	aaacgttccc	ctcaaggcaa	aaacaccctt	aagatgtatt	ctggagaatt	28800
gggaccaatt	tgactctcag	atgctaagaa	aaaaaaagaca	tattcttctg	cagtaccgcc	28860
tggaacgat	atactcttta	agggggagaa	acctggcatc	ctgaggggaag	cataaattat	28920
aacaccatct	tacagctaga	cctcttttgt	agaaaagaag	gcaaatgggtg	tgaagtgtca	28980
tacgtacaaa	ctttcttttc	attaagagac	aactcgcaat	tatgtaaaaa	gtgtgattta	29040
tgccctacag	gaagccctca	gagtctacct	ccctacccca	gcattcccca	gactccttcc	29100
ccaaataata	aggacccccc	ttcaacccaa	acggtccaaa	aggagataga	caaaggggta	29160
aacaactaac	caaagaatgc	caatattccc	cgattatgcc	ccctccaagc	gggtgggagga	29220
gaattcggcc	cagccagagt	gcacgtacct	ttttctctct	cagactttta	attaaaatag	29280
acctaggtaa	attctcagat	aaccctaattg	gctatatatga	tgttttacaa	ggtttaggac	29340
aatcctttga	tctgatattg	agagatatata	tgttactgct	aaatcagaca	ctaaccacca	29400
atgacagaag	tgtcgcgcta	actgcagcct	gagagtttgg	cgatctctgg	tatctcagtc	29460
aggtcaatga	taggtcgaca	acagaggaaa	gagaacgatt	ccccacaggc	cagcaggcag	29520
ttcccagtgt	agaccctcac	tgggacacag	aatcagaaca	tggagattgg	tgccgcagac	29580
atgtgctaac	ttgcgtgcta	gaaggactaa	ggaaaactag	aaagaagcct	gtgagttatt	29640
caatgatgtc	cactataaca	cagggaaagg	aagaaaatcc	taccgccttt	ctggagtgc	29700
taacggaggc	attgaggaag	catacctctc	tctgtcaact	gactctactg	aaggccaact	29760
aatcttaaa	gataagttta	tcaactcagtc	agctacagac	attaggaaaa	aacttcaaaa	29820
gtctgcctta	ggcccgaac	aaaacttaga	aaccctattg	aacttggcaa	cctcagtttt	29880
ttataataga	gatcaggatg	agcaggcaga	atgggacaaa	tgggataaaa	aaaaggccac	29940
cgcttttagtc	atggccctca	ggcaagcgga	ctttggaggc	actggaaaag	ggaaaagcta	30000
ggcaaatcaa	atgcctaata	gggttttgctt	ccagtgcggt	ctacaaggac	actttaaaaa	30060
agattgtcca	aatagaaata	agccgcccc	tcgtccatgc	acctcgtgtc	aagggaatca	30120
ctgtaaggcc	cactgcccc	ggggacgtag	gtcctctgag	tcagaagcca	ctaaccagat	30180
gatccagcag	caggactgag	agtgcgcggg	gcaagcacca	gcccattgcca	tcaccctcac	30240
agagccctgg	gtatgcttga	ccattgacgg	ccaggaggct	aactgtctcc	tggaactgg	30300
tgtggccttc	tcagtcttat	tttctgttcc	cagacaacgg	tcctccagag	ctgtcactat	30360
ccaaggggtc	ctaggacagc	cagtcaactag	atacttctcc	cagccactaa	gttgtgactg	30420
gggaacttca	ctcttttcac	atgcttttct	aattatgcct	gaaagcccaa	ctcccttgtt	30480
agggagagac	attctagcaa	aagcaggggg	cattatacac	ctgaacatag	gagaacaccc	30540
gtttgttgct	ccctgcttga	ggaaggaaatt	aatcttgaag	actgggcaac	agaaggacaa	30600
tatggacgag	caaagaatgc	ccgtcctgtt	caagttaaac	taaaggattc	tgctctcttt	30660
ccccaccaa	ggcagtaccc	ccttagaccc	gaggctcaac	aaggactcca	aaagattaag	30720
gacctaaaag	ccaaggcct	agtaaaagca	tgcaatagcc	cctacaataa	tccaacttta	30780
ggagtacaga	aaccagtggt	acagtggagg	ttagtgcaag	atctcaggat	tatcaatgag	30840
gtcactgtcc	ctctatacct	agctgtacct	aacccttata	ttctgctttc	ccaaataacca	30900

gaggaagcag	agtggtttac	agacctggac	cttaaggatg	cctttttctg	catccctgta	30960
catcctgact	ctcaattctt	atttgccttt	gaagatcctt	caaaccat	gtctcaactc	31020
acctggactg	tttcacccca	aggggttcagg	gatagccccc	atctattttg	ccaggcatta	31080
gcccagact	tgagccggtt	ctcatacctg	ggcactcttg	tcctttggta	tgtggatgat	31140
ttttactttt	agccgccagt	tcagaaacct	tgtgccatca	agtcacccaa	gtgctcttaa	31200
attttctcgc	tacctgtggc	tacaagggtt	ccaaacccaa	ggctcagctc	tgctcacagc	31260
aggttaaata	cttagggcta	aaattatcca	aaggcaccag	ggccctcagt	gcctattctg	31320
gcttatcctc	atcccaaaac	cctaaagcaa	ctaaggagat	tccttgacat	aacagggttc	31380
tgccaaatat	ggattcccag	gtacggcgaa	atagccagac	cattatatac	actaattaag	31440
gaaactcaga	aagccaatac	ccatttagta	agatggacac	ctgaagcaga	agcggctttc	31500
caggccctaa	agaaggccct	aacccaagcc	ccagtgttta	gcttgccaac	ggggcaagac	31560
ttttctttac	atgtcacaga	aaaaaacaga	aatagctcta	ggagtcctta	cacagggtcga	31620
tgagcttgca	acccatggca	tacctgagta	aggaaattga	tgtagtggca	aagggttggc	31680
ctcattgttt	atgggtagt	gcggcagtag	cagtcttagt	atctgaagca	gttaaaataa	31740
tacaaggaag	agatctgtgt	agacatctca	taacgtgaac	ggcatactca	ctgctaaagg	31800
agacttggtg	ctgtcagaca	accgtgagga	aagtaactaa	aatcgtaaat	ccccatggcc	31860
ctcccttatc	atatttttct	ctttactgtt	ctcttaccct	ctttcactct	cactgcaccc	31920
cctccatgct	gctgtacaac	cagcagctcc	ccttaccag	agtttctatg	aagaatgcgg	31980
cttcccagaa	atattgatgc	cccatcaaat	aggatgttac	ctaaaggaaa	ctccaccttc	32040
actgccacaca	cccatatgcc	ccacaactgc	tataactctg	ccactctttg	catgcatgca	32100
aatactcatt	attggacagg	gaaaatgatt	aatcctagtt	gtcctggaag	acttgagacc	32160
actgtctgtc	ggacttactt	caccataact	ggtatgtctg	aggggggtg	agttcaagat	32220
caggcaagag	aaaaacatgt	aaaggaagta	acctcccaac	tgaccgggt	acatagcacc	32280
cctagcccct	acaaaggact	agatctctta	aaactacatg	aaacctcca	tacctact	32340
tgcttggtta	gcctatttaa	taccaccttc	actgggtccc	atgaggtctc	ggcccaaaac	32400
cctactaact	gttggtatgt	cctccccctg	tatttcaggc	catgcatttc	aatccctgta	32460
cctgaacaat	ggaacaacta	cagcacagaa	ataaacacca	cttcctgttt	agtaggacct	32520
cttggtttcca	atctggaaat	aaccataacc	tcaaacctca	cctgtgtaaa	atthagcaat	32580
actgtagaca	caaccaactc	ccaatgcac	aggtgggtta	ctcctccac	acgaatagtc	32640
tgccctacct	caggaatatt	ttttgtctgt	ggtaccttag	cctatcgttg	tttgaatggc	32700
ttctcagaat	ctatgtgctt	cctctcatte	ttagtcccc	catgaccatt	tacactgaac	32760
aagatttata	caattatggt	gtacctaagc	cccacaacaa	aagagtactc	attcttcctt	32820
ttgttatcgg	agcaggagt	ctaggtggac	taggttctgg	cattggcggt	accacaacct	32880
ctactcagtt	ctactacaaa	ctatctcaag	aactcaatgg	tgacatggaa	tggttgccg	32940
actccctggt	caccttgcaa	gatcaactta	acttcttagc	atcagtatgc	cttcaaaatt	33000
gaagagcttt	agacttgcta	acctctgaaa	gagggggaag	ctgtttattt	ttagggaag	33060
aatgttggtta	ttatgttatt	ttagcgggaag	aatgttggtta	ttatgttaat	caatcctgaa	33120
ttgtcacaga	gaaagtga	gaaattcgag	attgaataca	acgtagaaca	gaggagcttc	33180
aaaaacacca	gaccctgggg	cctcctcagc	caatggatgc	cctggattct	ccccttctta	33240
ggatctctag	cagctcta	attgatactc	ctctttggac	cctgtatctt	taacctcctt	33300
gttaagtttg	tctcttcag	aatcaaagtt	gtaaagctac	aaatcgttct	tcaaattgaa	33360
ccccagatga	agtccatgac	taagatctac	cgtggacccc	tggaocggcc	tactagccca	33420
tgctccaatt	gtaatgat	cgaacgcacc	cctcccgagg	aaatctcaac	tgcaaacacc	33480
ctactatgcc	ccaattccgc	aggaagcagt	tagactggtc	gtcagccaac	ctccccaaca	33540
gcacttggtg	tttcctgttg	agtgggggga	ctgagagaca	ggattagctg	gatttcctag	33600
gccgactaag	aatcccaaag	cctagctggg	aaggtgacca	catccacctt	taaacactgg	33660
gcttgcaact	tagctcacac	ccgaccaatc	aggtagttaa	gagagctcac	taaaatgcta	33720
attagacaaa	aacaggaggt	aaaaaaatag	ccaatcatct	atcgctgag	agcacagcgg	33780
gaaggacaat	gatcgggata	taaacccagg	cattcaagcc	ggcaacggct	accttctttg	33840
ggtccctcc	ctttgtatgg	gagctctctc	tgtcttca	ctattaaata	ttgcaactgc	33900
aaaaaaaaaa	tagcttaatt	gaagaataaa	ttaatacaat	aaaaggaata	cattttaagt	33960
atacagttca	aactgtaaca	gtgttacagt	ttcaagagga	ccccttcaac	aagatattgg	34020
gcatttccat	catgccttaa	aagttccttc	ttgtccctta	ctgggtgggt	ccatctctac	34080
tacacctctc	tgacctggcc	cagaccttgg	cctcagaaga	atcatttttt	tgctactaca	34140
tattagtttt	gtctgttcta	gaacttctta	aaaacagaat	catagagtat	gttctctttg	34200
tattggttct	ttttactcaa	tgtaatgttc	tgtgacattt	atccatatta	ttgcatgtat	34260
tattcctttt	aatcctgaat	agtatgctgt	tttaggaata	taatgcaatt	gtttattcat	34320
ttacctgttg	acagatatct	gagctattat	gatggatatt	atgaataatt	ctgctatgaa	34380

cactttctgta	caatgttttc	tcggacatat	attttccattt	ttcttgagt	gagctgttag	34440
aactgttgga	tcagaaagta	agcatatgtt	gaattttgaa	agaaactgg	aaactctgt	34500
ctaaagtgat	ttgtaccatt	ttacactcct	actaataatg	tatgagagtt	atatttgcct	34560
cacagccttt	ttactacttt	gttaatcttt	ttagtactgt	caaccttttt	aatttatcca	34620
atctagggaa	cgtgaagtag	tatctcactg	ttattttcat	tttctgatg	agtaacaata	34680
tcgtgtatct	tttcatgtgc	ttattagcca	ttcctatatc	ttttgtgaaa	tagttaactt	34740
aaatttgtaa	ctaaagggtg	tttcttgagt	ttcaggtagt	aagcctattt	ccctcaagtg	34800
aataaactac	agtcttgga	tgaaaaatta	aacacagtg	agacattttt	tgtataagtt	34860
gttttactct	gtgtatgtct	ggtttgctta	gtctattatt	atatgcccc	tgaaagcaaa	34920
cacagtgtct	atctcactaa	tgagtatcac	tagcacatag	aactgtgctt	gccccaaagc	34980
tgaactcaat	aaatatgtta	atgtgtatgc	atgcacatac	atctacatgc	atgtacatct	35040
atacacacac	ataaacatat	attaattttt	agaccacaaa	atctaagaaa	actaattctt	35100
gagcctctgg	tttgaagaat	tctcaaat	tttatgttcc	actccacatc	actccacatc	35160
cactgtacct	gaaatagccc	tactgttcta	cttttggtaaa	tcaggcaaat	tttaattttt	35220
aaataattaa	gattccaact	aattttaaaa	tataatttga	aagttaacaa	tgaaatacat	35280
tacataaaaa	gaaaatttta	aataaaagca	aaactaaacc	caataagagg	aaagaaagtt	35340
gggctgtatt	tctttaatcc	tttaaaattc	aaatcacaca	atgctccaat	gaaatcttca	35400
ttaaactgaac	caaactatgc	ccatgaaaga	tctcatatgc	aactgctaaa	acctcaataa	35460
acataattcat	cttcttgcaa	aaaagatatt	tctttataat	atgcacatgc	agtatatact	35520
attttgaggc	agatttgtac	tttagtctct	gttccattgc	ttaccggctg	gctgtccttt	35580
gtctgggtcat	tgacctccaa	cttaaaaaat	aatacttgcc	ttgtctacct	cacagaagtg	35640
ttatgaaagt	caaacaaggt	agcataaagg	tatttttaca	gatataaagt	gctataatac	35700
agatttttaa	aatcactcta	catcccataa	tactttgttg	tacaatttta	gagcaatagt	35760
agaaaataac	aattattggc	taattgaaaa	tccagtcctg	aattccataa	aatgtatgat	35820
atgaacatta	tagtacatca	tattacgagc	cccaaataat	cactgcttat	atagttggtt	35880
aggatttctt	tagtttggtc	atatagttta	tatatattatg	cagtccctat	tttgtgagag	35940
gcattgtgag	gagcataaag	acataagcac	agtacagagc	cttagcttct	ctacattttac	36000
taaagaagac	ttcttcttgg	gtattttaatc	aattattttaa	gtatttctggg	aagaaatgaa	36060
attaacttca	tagactgacc	ttagattact	atcattacaa	aaagatgcct	gagtgtactg	36120
tctttaacat	accagtat	atcttataac	tgttatattt	acttgaatca	gaagtgaagt	36180
ccttttaagc	actaagcatc	cattctatac	tttcttgtct	ttacatatga	gatacaaatc	36240
atatttttaa	aactttttatt	tactttttatt	tttttagagac	ggagtcttgc	tctgtagccc	36300
aggctggagt	acagtggcat	gatcttggct	caccacaatc	tccacctcca	cttcccaggg	36360
ttcaagtga	caaatacatc	ttttaagcac	agattctcaa	catgtatcct	agcatgctac	36420
tgccataact	agggtgtgaa	ttaaagtatta	aagacagctt	accccaaata	ttactgtaac	36480
atatactctc	aaatgaaaaa	gaacatatta	acaactatac	ttggatggga	ttctgggagc	36540
taaccatcc	ctctctcccc	tttcttccaa	attccatctc	ctattaacac	accagctctc	36600
ctgagctaag	cagctcctgg	gggtggggaa	gggtgtacat	ggagaaagct	agaacctcta	36660
cagtgttttc	ctctctggga	ggaactagca	ggcatacgaa	cagaaaaagc	tgaataaaag	36720
gctgaatcct	ttctattcct	gaggcagaca	gagagaagac	cagggaaaca	agagacttcg	36780
accaagagcc	ctgccaggta	ttgatacctt	tgatactgag	aaaatatctg	ggatatgaaa	36840
tacaaatgct	aaataagtat	ctttgaaata	ggggtaaaag	aataaagggt	cttgatgagt	36900
aaaatgggta	gtatttttta	ataacctgat	aatgagcttt	aggaaaagg	aagggtcaacg	36960
ttatggaatg	aaaacacaga	ggtaccaa	ttaaaagcat	aaaaaaaagt	ggaggggggg	37020
aaccataata	cttcatcaaa	ctagcaataa	acttagtatc	atttctaatt	agaaacgcta	37080
gaaggaaatc	acttagatct	gataaagact	aggctataat	tctaactgat	gaaacactta	37140
aactgtatca	attaatacca	gaaaacaaac	acagaaaagt	ctactagaac	catcattatt	37200
cagcacagtc	ttggtaatgc	aatactataa	tagcaatgca	ataaagcaag	aaaaaaaaaa	37260
gtttgtaaaa	acacaatagg	atgagatttt	tgtttttcca	atgccataaa	taactagaaa	37320
tggaaacaaa	ataaagaaaa	acaaaatcta	caaaacacct	ggaaataaaa	agaaaaatgg	37380
tctatttgaa	gaaaacctta	aaatctatgc	agaacataaa	acaaaatctg	aataaaaaag	37440
aatatcatgt	tcttgtctgg	gaagacttaa	tatcataaga	aagtgaatta	tatcaaaatt	37500
taaaatcgaaa	tttaatgtat	ttccatctct	aatcagacag	gacactatgg	ggaactgaat	37560
aagtgtattt	aaaagtcag	gaaaattaat	aactgagaat	aacctgaaaa	agtatgaaaa	37620
aaggagacaa	atgaattgct	ccaacagata	tcagaacgct	aaaattaaat	aaaaatacta	37680
ctaggataag	aaaatacata	tactgatgta	atgaataaag	aatccagaat	tagattccag	37740
taagtcaaac	tactttacta	taaaccagg	gtggcatatt	catccagtg	gaaaaggaca	37800
gtaagaagtg	agtaaaactat	ggccactgg	ccaaattgtg	gcctctgcct	atttttgcaa	37860

ataaagtttt	actgggacaa	agccaagcct	atcatttgca	aattgtctat	aaatatatttc	37920
atgttacaga	atcacacagt	ttcaacagag	accatcttgt	ctacaaagct	gaaaatatct	37980
actatctggc	ccttgaagaa	agtttgccaa	accttagttt	atataataaa	agatcagcta	38040
tctcatagac	acctatctca	cacaacacat	tgtgggaaa	gaccttcttt	tttttttgag	38100
acggggtctt	gctctgttga	ccaggctgga	ctgtagtggc	atgatcatgg	ctcactgcag	38160
cctcaacctc	ccaggttcaa	gtaatgctcc	caccacagaa	tcccaaacag	ctgggagaga	38220
tgtgtgccac	tacgcctggc	taaggggcct	ttttaacaga	gaaagaaatc	cacatactac	38280
taagaaaaag	aagggcata	ttgatata	tttatatttt	ttatatagat	atcataaaaa	38340
tcaagatgaa	ttatacagtt	atattttgca	atgtgtttga	cggtaaaagt	ttaatatcta	38400
taaaaattat	tttataaaat	atctttaata	tatttataga	tattataata	taaaatatct	38460
ataaaattat	tttataaaat	aaaaagttaa	gaagaaaaga	taggcaaaac	aaaatacagt	38520
gcaatttaca	gaaaaccaag	tccaaatggt	caacaaagat	aaaacagatt	tataaactca	38580
ctaagtgtga	gagaattatt	agttaaaagta	aaaatatctc	tctataccca	caatactact	38640
aaaaatcaga	gttataatgc	cctattgctg	gtggagatgt	aaggggagaa	gcargctctc	38700
atatactgtt	agtgaataat	taaactaata	catttttgaa	aagtaagctg	gcaatttttt	38760
ttttaatctc	taccttttga	tgcaaaaact	catttttggt	tacctattcc	ataccttaaa	38820
aaaaatacat	atgcttactg	tagtactgtt	tataatggta	aaaactagaa	aaaaagaaaa	38880
cttgatagt	aatactgaac	aaattacag	gcactacag	attaaacata	atgcagccat	38940
taaaaagaa	taaattaggc	tggtgctggt	ggctcatgcc	cgtaatccca	gcactttggg	39000
aggccaaagc	aggcggatca	cttgaggcca	ggagttcgag	accagcctgg	ccaacatggc	39060
aaaaccctgg	ctctacaaaa	aatacaaaaa	ttagtcgggc	atggtggtgg	gcacctgtag	39120
tcccagctac	tcaggaggct	gaggcaggag	aatcacttga	gcctgggaga	cagagattgc	39180
agtgaagcaa	gatcatgcca	cagcattcca	gtccagggtga	cagaacgaga	ctctgtctca	39240
acaaaagaa	caaattaaac	cctacaactc	atcaacaaaa	atacccaaac	ccaattcaaa	39300
aatgggcaaa	ggacttgaat	agacatttct	tcaaggatga	taaacaagca	catgaaaaga	39360
tgacagcac	tattcattag	tgattacatc	ccacatgcat	taggatggct	agtatgaaga	39420
acagaaaata	ataaatattg	gtgaagatct	gaaaaacaga	aacctttgtg	caactgttgt	39480
gggaatgtaa	agtgggtacag	ctactacgga	aaacagtatg	gccattcctc	aagaaaataa	39540
aaataaaaatt	atcttatgat	aggaatatgc	atttctgggt	aaatacccca	aataactgaa	39600
aacagggtgt	acacccattt	caacatttac	atgtcaattc	aactgggcca	gaataccag	39660
atatattgttc	aaatattctt	ctggatgctt	ctatatatat	gttttttggc	tgagggttaac	39720
attttaaattg	gtggattctg	agtacagcag	attaccatcc	acaatgtagg	tgggcctcat	39780
ctactcagtt	gaagggtctta	cagaaaaaga	ctgacctccc	ttgagcaaga	aagaattcag	39840
gcaacagact	gccttttgac	tcaactgcaa	ctcttccttg	agtcaacagc	ccatcccac	39900
acccctggctt	ggtgagtcca	gggtctgatg	aggtaggctg	cagactcaag	gaagagctgc	39960
cacacacagc	aaagccaatt	cattaaaaata	aatctctctc	tacacaaaaca	caacacacaca	40020
ctaccaccac	caccatgatg	gttctgtttc	tctggagaat	gctaatacac	ccctgttcat	40080
ggcagcatta	ttcacaaatag	ccaaaagggtg	gaagcaactc	cagcagatga	atggagaagc	40140
aaaatgtggt	atgtatatatac	aatggaatat	tattaagcct	ttaaaaagt	gaaatttatat	40200
ctatctatat	ctatacacac	atactcacac	acacacacac	acatttatag	aagacagggt	40260
ttcaccatgt	tgtcaaggct	ggtctcgaac	tcttgggctc	aagcaaacccg	cctgcctcag	40320
cttcccaaa	tgctgagatt	acatgtgtga	gccaccacac	ccagccaaaa	aaaggacatt	40380
ctgacacata	atacaatata	gataaacaat	gaggacatca	tgatatgcga	aataagcctg	40440
tcacaaaaag	gcaattagtg	tatgattcct	cttgatgag	gtacctatgg	atgtcaaate	40500
cataaagtag	aatggggaaa	cagagagttg	tttaatgggt	atagagtttg	ttttgcaaga	40560
agaaaaagagt	tttgaggaaat	gaatgtacaa	cagtgtgaac	ataattaaca	ctactgaaaa	40620
tggttaagat	tataaatttt	atgttacatt	tattttacca	tgattaaaaa	ttaaaacaaa	40680
ataatattaa	ggaaaaatac	tataaataac	aacaacaaaa	aaaacacctc	aagcaactta	40740
cattcacctg	ggaaacagaa	tacatcctat	tctgctagag	atataatctgc	agttcaaaat	40800
ttattacaaa	tgatgttgtg	tatctttttg	aaatgactga	aaaactaaat	taaaagcaat	40860
aatattcagt	ttactaacca	gtaagtcctt	ctttcatggt	tcctgacttt	tctgtaagat	40920
gttattgcaa	gatatctact	aaaatggaaa	acaactgaaa	aggcaaaatt	ataattttct	40980
atcaacatcg	ctaaaaccct	ggaggggaag	aactctaaca	aacatggcca	taatttgcca	41040
catatttcta	ctgtcctcac	ttttcaaaat	ccagaaatca	acattttctg	aaacaaaaca	41100
gagtctaaaa	tttggtcctc	tcttcagttt	agaagggtgcc	aagttaatcc	ctgacatcct	41160
agtttccatt	ttcaaaaatg	tactttttct	ctcccaaac	cggtatctag	attcttaaat	41220
atttttagca	catagaagtt	aaatagattt	gcttaaccaa	aatagccagt	aaacctccca	41280
aaagaattaa	aatattaatg	gcgctttaat	gatacaaatg	aacaacttta	cattcaatcg	41340

tcaatgggaa	aggaagcaga	attctgagga	ttaatgaaagt	aaacaaaacg	aagttcaaat	41400
tctactttat	tttacttttt	tgtaactaat	gaacaacttc	ttccaaagac	aagtaggaaa	41460
tacaaaaatt	agccaggcat	ggcacatgcc	tgtagtcctg	gttacttgga	aggctgaagt	41520
gggtggatcg	cttgagccgg	gaaggcagag	gctgtagtga	gctgagatca	catcac-gca	41580
ctcaagcctg	ggtgacagag	caagaccctc	tctggggaaa	aaaaaaaaaa	aaataggctg	41640
ggcgagctgg	ctcacacttg	taattccagc	actttgggag	gctgaggcag	gtggttcacc	41700
tgaggtcagg	agttctagac	cagcctgacc	aatatggtga	aacctgtct	ctactaaaaa	41760
tacaaaaatt	agccaggcat	ggtggtgggc	aattgtaatc	ctagctactc	gggaggctga	41820
ggcaggaaaa	tcgcctgaac	ccaagaggcg	gaggtttcag	tgagccgaga	ttgcactagt	41880
gcactccagc	ctgggcgaca	gagcaagact	tcatctcaaa	ataaataaat	aagtaagtaa	41940
ataaaattaa	aaaatatata	aaaataaaac	aaagataagt	aggaaccatc	cttttttttt	42000
tttttttttt	ttttttttta	agatagggtc	tgtttctgat	gcccaggctt	gagtgtagtg	42060
gcctgatcat	ggctcactgc	aaccttgacc	tctcaaatac	aagtgactct	cctacctcag	42120
cctcccaagt	agctgggact	acaggtgctt	accaccccat	ccggctcatt	taaaaaaat-	42180
tttttgtaga	ggtggggtct	cactatggtg	tatccaggct	ggtctcattt	taactttat-	42240
agaaaacaag	cattgtttta	tcagcttctt	gtttttttta	aactaaaaat	aacactgcta	42300
ggttgtttct	atgaagattc	tctaaattta	tttataacct	taagaataac	atgtagaaca	42360
aagtagatga	ctgaatgatc	tttggtgaat	aaatatgaat	ggatattcaa	ataattaaaa	42420
atctcttaag	atctcccat	ctttacagga	tacagagaaa	actcgttaat	atggcctgac	42480
ttttaccttt	gcagccttat	ccaaactctg	tggtcaagac	aaacaggttg	tccttatact	42540
tacaacgtcc	ccctttgcct	acaaagctct	tctcatgact	ctttgcctat	cttaagttca	42600
cctatctgtc	aaatctctgg	gaatgcaaca	tttctcctca	gtagccttct	ctcctcccaa	42660
actagaacaa	attcttctctg	gggcatttagg	tttttattgc	actgtatgtc	tcttcttcac	42720
agcaatcaca	gttccaatgt	tatatttgta	ttcttagttg	atgtgtttct	ttccaccttt	42780
agactataac	cttctaaggg	gtcacacata	atatcgatca	tcagttgtat	cccttgtgca	42840
tagcacaggg	catggcaggc	aaatatgtgt	gtaaataaac	ttggtgaatg	aatcaatgag	42900
acacactttt	cttaccctaaa	gtataatggc	aggataacat	ttatcaatct	attgcttctt	42960
gaaaaacaga	tatgatgtgc	ttaattttca	ttttacatct	caaataccaa	tgccctaagga	43020
attcacagtc	attttacaaa	tctttttgac	aaatgccttc	attaatcacc	acctgtttac	43080
aagtgtctaaa	taacattttg	gttacattct	gtaacatttc	ctgcacttaa	tgtcactctc	43140
agaatactgg	ctaatatgaa	gcacctggac	ttcaggaaca	caaacctgaa	actaacacac	43200
caaactaaac	tgttatgtaa	atgacagaaa	tgacacattt	tggtctgcaa	catctctaga	43260
tggtctttgg	accaattcaa	cttttaccac	taaaaaatcg	tcacctgact	atagtcattt	43320
tgagctcatg	ataaatgaat	tacagatgaa	aaataaatag	tttgatgaca	atcttttaca	43380
aagtttatct	tcaaagaata	ccaccagtca	caggtattct	aggctcctat	caactttatt	43440
ggtcagggtc	gacttcactt	ttcatgataa	ttatgttctg	aaaattctac	aaacttaatg	43500
attacaaaca	aaagtcatag	tttgctcata	aatcaggcct	aggctctggat	tctagttctt	43560
ccatttttca	tttgttcact	gaggcaagtg	acttaaaatt	ccctagcctc	agtttctctc	43620
catgtaaaat	cagataatga	ttcctattcc	taagatgggt	ttgaggcttc	aacaagataa	43680
gatgggcctc	actcaagcat	gctcagtact	ctgtctctct	ctctccggtt	atgcagaaat	43740
tctattagga	ttctgcaaag	taaaataaat	atttcagtaa	aaatttatgcc	ctttattaat	43800
gaatctagat	tttcagattt	tccttaaat	tacttagtaa	cttaagggtc	caaataattat	43860
agagatttgt	atctagtatt	ttaaagaaat	gaaagggtgt	aatcaaaatg	ctgcacaaat	43920
aaatgtaca	tttaacaaac	agaatatcac	aaccatacaa	actaatcaga	tataaagaag	43980
tcagcaacag	aaatctgatg	ttgcctttag	atcacacaat	taggcaaaca	aaaatagagt	44040
tccatcctcc	tttggtcaag	gccatgggtg	aagactgaat	accaaataag	gaaataggaa	44100
aagccaggaa	atggcaaatt	agcaaaaact	ggactcctta	atttttatat	tcattttcat	44160
atctcacttc	taaaaacttta	attaaattca	aataaaaaacc	aaaatggaac	tgagataaag	44220
ccaaaaggaa	agttatgtag	gtcaaatgag	aacctatatt	gtccttaggc	tctttgttgc	44280
tttctgttta	aggaaaaact	gcccaagtgc	cttgacacat	taaagatcaa	gcaggagggt	44340
ctgccgagag	tccccatctg	gcagccaggt	tttgtcaagc	aaattttgag	aattctctac	44400
cctcccactt	tctatctaat	tatagcactt	tataaaaaacc	attctctctc	tgtctctgtc	44460
tctctctctc	tctctctctc	acacacacac	acacacacac	acacacacac	acacacaccc	44520
tttctctctc	tctctctctg	aaacttatct	gtattataat	aacacaacac	taggtatgga	44580
ttaatctgac	aattttcccc	taaaacagaa	taaattcaaa	aaggaaaacc	tttctctgtt	44640
acacatgcac	tatattctga	caataataat	tcctaaatta	agtataatac	attttcccta	44700
caggagttaa	aagaagttac	agtaaagaat	ctcttgataa	aatatatatg	ccagaacttg	44760
acccaaataa	gtgctgagag	gtataaatct	caaaacagtt	tccggactct	ttgtgaaatg	44820

tcttcagagt	ctgcgatata	ttttcttcaa	ctaaattata	caagtaagat	attttgcctgg	44880
gctgtgggaa	tgccttacgg	catgttactg	tggagctcat	ggtaaaatag	aaagaatata	44940
aataattaaa	ataaaattga	caaattgataa	atgatttaaat	aaattagaaa	ttcaaattgcc	45000
gggcactttt	ctagaacctg	gacacaaagc	atgaacctaa	caataacccc	gccttcatga	45060
aaaaatagga	ctattttgaa	attatacctg	caacactaaa	taaatattct	tcatctctcc	45120
agtatattga	gatgttttact	ttcaattaga	caatttgcct	tcctctctga	acacatagtt	45180
atgtgatggc	tctataaaaag	attttataaa	aactatagaa	ggaactattg	gtaaaagactg	45240
tgggatacta	aaaatggcta	caaagaaagt	tatgacaaaa	cctctgagtt	tgaatggaag	45300
tcctactaga	ttagagtcta	agcctgtgac	attatgcttc	tggttcttgt	tcttaaatgc	45360
ttttctcatt	aatagtatgt	aacttacttc	ctggaatgcc	attcattaaa	aaaatatttta	45420
atattttgcta	aatgtcaata	ttttgcccag	cactttttaa	gtacagaaaac	atggagtttc	45480
ttttacctcat	gcaaattatgc	tgtgagaaga	acttaagagc	ctattgccta	ctttgtggta	45540
caacactgaa	gactcaccat	ccaaaaaaaa	cagacttagt	aaattcttgt	gatttgcagt	45600
agttctgttc	tataaggtta	ccacaaacac	tgaaatcatc	gctcctgggg	gaatacaagg	45660
ttatgtttcc	gtgagccctc	ggtcacacaa	tgttcattaa	ctgatcaata	cataaccttg	45720
ttctatgtgt	gtttctgttt	aaaaagagca	cttcagtgct	acatttgagg	tctgttttaa	45780
acagcaaaat	cactaataaa	aagcacaaaa	atgtaaaagc	atggcactac	atacactgtg	45840
acaagaaggc	ttgtttatag	tatgacagct	gagacaagaa	ggtagagcct	cgctttgatc	45900
aacctctgct	gggaaatgag	catcagggtga	atcaattttt	caccactctg	aatgaccgta	45960
aaagtgtctc	aagtactgac	tttgggggtta	cacataaatt	ttagtaagca	tgtgaatctg	46020
ccaatatgaa	atctacaaat	aatgagtacc	aatgcatat	gagtcaaata	tttcagtgcg	46080
gtatctgact	tgattgccac	tgaaagacac	agtttggaag	accctaata	aataccgttt	46140
agttactatg	cagacaaaaga	gtttctacact	agagtgcttc	aattaagatg	tctgaggctt	46200
tcataaatgg	atgttttttta	aaatgtttatt	tcctacctga	tatattctaa	aggggatata	46260
acgaaatcca	ttttcttctg	caggatattc	catgagtttc	cgattgatgg	cccaaaactg	46320
gtcaaatctg	tctgtaatga					46380

```
<210> 67
<211> 773
<212> ADN
<213> Homo sapiens
```

<400> 67						
actgagagac	aggactagct	ggatttccta	ggctgactaa	gaatccctaa	gcctagctgg	60
gaaggtgacc	acatccacct	ttaaaccacg	ggcttgcaac	ttagctcaca	cctgaccaag	120
gaaggtgacc	acacctcct	ttaaaccacg	agcttgtaac	tcagctcaca	cccgaccaat	180
caggtagtaa	agagagctca	ttaaataacc	aattaggcta	aaaacaggag	gtaaagaaat	240
aatcaaataca	tctatcgct	gagagcacag	ggggaggggac	aatgatcggg	atataaaacc	300
aggcatttga	gccagatcag	gtaacctct	ttgggtcccc	tcacactgta	tgggagctct	360
gttttctactc	tattaaatct	tgcacctgca	cactcttctg	gtccatgttt	gttcggctc	420
aagctgagct	tttgctcgcc	gtccaccact	gctgaatgcc	gccattgcag	acctgccctt	480
gacttccacc	cctccgcatc	cggcagagtg	tccgctgcac	tcctgatcca	gcgaggcacc	540
cattgccact	cccgatcagg	ctaaaggctt	gccattgttc	ctgcacagct	aagtgcctgg	600
gttcatccta	atcaggctga	acactggctg	ctgggttcca	cggtttctct	ccatgactca	660
cagcttctaa	tagagctata	acactcacca	catggcccaa	ggttccattc	gttgaatcc	720
atgaggccaa	gaacccagg	tcagagaata	aaaggcccg	cccatcttgg	gag	773

```
<210> 68
<211> 10
<212> PRT
<213> Homo sapiens
```

```
<400> 68
Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val
  1                   5                   10
```

73

<210> 69
<211> 10
<212> PRT
<213> Homo sapiens

<400> 69
Leu Leu Phe Gly Pro Cys Ile Phe Asn Leu
1 5 10

<210> 70
<211> 10
<212> PRT
<213> Homo sapiens

<400> 70
Cys Leu Pro Leu Asn Phe Arg Pro Tyr Val
1 5 10

<210> 71
<211> 10
<212> PRT
<213> Homo sapiens

<400> 71
Gly Leu Leu Ser Gln Trp Met Pro Trp Ile
1 5 10

<210> 72
<211> 9
<212> PRT
<213> Homo sapiens

<400> 72
Cys Leu Pro Ser Gly Ile Phe Phe Val
1 5

<210> 73
<211> 9
<212> PRT
<213> Homo sapiens

<400> 73
Trp Met Pro Trp Ile Leu Pro Phe Leu
1 5

74

<210> 74
<211> 10
<212> PRT
<213> Homo sapiens

<400> 74
Ile Arg Trp Val Thr Pro Pro Thr Gln Ile
1 5 10

<210> 75
<211> 10
<212> PRT
<213> Homo sapiens

<400> 75
Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu
1 5 10

<210> 76
<211> 10
<212> PRT
<213> Homo sapiens

<400> 76
Leu Arg Thr His Thr Arg Leu Val Ser Leu
1 5 10

<210> 77
<211> 10
<212> PRT
<213> Homo sapiens

<400> 77
Lys Arg Val Pro Ile Leu Pro Phe Val Ile
1 5 10

<210> 78
<211> 10
<212> PRT
<213> Homo sapiens

<400> 78
Cys Arg Cys Met Thr Ser Ser Ser Pro Tyr
1 5 10

75

<210> 79
<211> 10
<212> PRT
<213> Homo sapiens

<400> 79
Thr Arg Val His Gly Thr Ser Ser Pro Tyr
1 5 10

<210> 80
<211> 10
<212> PRT
<213> Homo sapiens

<400> 80
Ala Arg Glu Lys His Val Lys Glu Val Ile
1 5 10

<210> 81
<211> 10
<212> PRT
<213> Homo sapiens

<400> 81
Ser Arg Ile Glu Ala Val Lys Leu Gln Met
1 5 10

<210> 82
<211> 10
<212> PRT
<213> Homo sapiens

<400> 82
Ser Gln Trp Met Pro Trp Ile Leu Pro Phe
1 5 10

<210> 83
<211> 9
<212> PRT
<213> Homo sapiens

<400> 83
Cys Tyr Tyr Val Asn Gln Ser Gly Ile
1 5

76

<210> 84
<211> 9
<212> PRT
<213> Homo sapiens

<400> 84
Phe Tyr Tyr Lys Leu Ser Gln Glu Leu
1 5

<210> 85
<211> 9
<212> PRT
<213> Homo sapiens

<400> 85
Thr Tyr Thr Thr Asn Ser Gln Cys Ile
1 5

<210> 86
<211> 9
<212> PRT
<213> Homo sapiens

<400> 86
Ser Phe Leu Val Pro Pro Met Thr Ile
1 5

<210> 87
<211> 9
<212> PRT
<213> Homo sapiens

<400> 87
Tyr Tyr Val Asn Gln Ser Gly Ile Val
1 5

<210> 88
<211> 9
<212> PRT
<213> Homo sapiens

<400> 88
Leu Phe Asn Thr Thr Leu Thr Gly Leu
1 5

77

<210> 89
<211> 9
<212> PRT
<213> Homo sapiens

<400> 89
Leu Phe Gly Pro Cys Ile Phe Asn Leu
1 5

<210> 90
<211> 9
<212> PRT
<213> Homo sapiens

<400> 90
Arg Trp Val Thr Pro Pro Thr Gln Ile
1 5

<210> 91
<211> 10
<212> PRT
<213> Homo sapiens

<400> 91
Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile
1 5 10

<210> 92
<211> 10
<212> PRT
<213> Homo sapiens

<400> 92
Leu Pro Tyr His Ile Phe Leu Phe Thr Val
1 5 10

<210> 93
<211> 10
<212> PRT
<213> Homo sapiens

<400> 93
Gly Ala Leu Gly Thr Gly Ile Gly Gly Ile
1 5 10

78

<210> 94
<211> 10
<212> PRT
<213> Homo sapiens

<400> 94
Leu Pro Phe Val Ile Gly Ala Gly Val Leu
1 5 10

<210> 95
<211> 9
<212> PRT
<213> Homo sapiens

<400> 95
Arg Arg Pro Leu Asp Arg Pro Ala Ser
1 5

<210> 96
<211> 9
<212> PRT
<213> Homo sapiens

<400> 96
Phe Arg Pro Tyr Val Ser Ile Pro Val
1 5

<210> 97
<211> 9
<212> PRT
<213> Homo sapiens

<400> 97
Arg Arg Ala Leu Asp Leu Leu Thr Ala
1 5

<210> 98
<211> 9
<212> PRT
<213> Homo sapiens

<400> 98
Trp Arg Met Gln Arg Pro Gly Asn Ile
1 5

79

<210> 99
<211> 10
<212> PRT
<213> Homo sapiens

<400> 99
Asp Arg Ile Gln Arg Arg Ala Glu Glu Leu
1 5 10

<210> 100
<211> 10
<212> PRT
<213> Homo sapiens

<400> 100
Leu Arg Thr His Thr Arg Leu Val Ser Leu
1 5 10

<210> 101
<211> 10
<212> PRT
<213> Homo sapiens

<400> 101
Glu Arg Val Ala Asp Ser Leu Val Thr Leu
1 5 10

<210> 102
<211> 10
<212> PRT
<213> Homo sapiens

<400> 102
Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu
1 5 10

<210> 103
<211> 10
<212> PRT
<213> Homo sapiens

<400> 103
Gln Phe Tyr Tyr Lys Leu Ser Gln Glu Leu
1 5 10

80

<210> 104
<211> 10
<212> PRT
<213> Homo sapiens

<400> 104
Gln Trp Met Pro Trp Ile Leu Pro Phe Leu
1 5 10

<210> 105
<211> 10
<212> PRT
<213> Homo sapiens

<400> 105
Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val
1 5 10

<210> 106
<211> 10
<212> PRT
<213> Homo sapiens

<400> 106
Asn Phe Val Ser Ser Arg Ile Glu Ala Val
1 5 10

<210> 107
<211> 9
<212> PRT
<213> Homo sapiens

<400> 107
Gly Pro Leu Val Ser Asn Leu Glu Ile
1 5

<210> 108
<211> 9
<212> PRT
<213> Homo sapiens

<400> 108
Leu Pro Leu Asn Phe Arg Pro Tyr Val
1 5

81

<210> 109
<211> 10
<212> PRT
<213> Homo sapiens

<400> 109
Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile
1 5 10

<210> 110
<211> 10
<212> PRT
<213> Homo sapiens

<400> 110
Glu Pro Lys Met Gln Ser Lys Thr Lys Ile
1 5 10

<210> 111
<211> 10
<212> PRT
<213> Homo sapiens

<400> 111
Leu Pro Tyr His Ile Phe Leu Phe Thr Val
1 5 10

<210> 112
<211> 9
<212> PRT
<213> Homo sapiens

<400> 112
Arg Glu Lys His Val Lys Glu Val Ile
1 5

<210> 113
<211> 10
<212> PRT
<213> Homo sapiens

<400> 113
Lys Pro Arg Asn Lys Arg Val Pro Ile Leu
1 5 10

82

<210> 114
<211> 9
<212> PRT
<213> Homo sapiens

<400> 114
Val Val Leu Gln Asn Arg Arg Ala Leu
1 5

<210> 115
<211> 10
<212> PRT
<213> Homo sapiens

<400> 115
Ala Val Val Leu Gln Asn Arg Arg Ala Leu
1 5 10

<210> 116
<211> 9
<212> PRT
<213> Homo sapiens

<400> 116
Leu Pro Phe Val Ile Gly Ala Gly Val
1 5

<210> 117
<211> 9
<212> PRT
<213> Homo sapiens

<400> 117
Asp Leu Tyr Ser Tyr Val Ile Ser Lys
1 5

<210> 118
<211> 10
<212> PRT
<213> Homo sapiens

<400> 118
Thr Glu Gln Asp Leu Tyr Ser Tyr Val Ile
1 5 10

<210> 119
 <211> 2615
 <212> ADN
 <213> Homo sapiens

<400> 119
 gaattccggg aagccagacg gttaacacag acaaagtgtc gccgtgacac tcggccctcc 60
 agtgttgcgg agaggcaaga gcagcgaccg cgcacctgtc cgcccggagc tgggacgcgc 120
 gcccgggcgg ccggacgaag cgaggaggga ccgcccaggc tgcccccagg tgtaactcca 180
 gcactgtgag gtttcaggga ttggcagagg ggaccaaggg gacatgaaaa tggacatgga 240
 ggatgcggtat atgactctgt ggacagaggc tgagtttgaa gagaagtgtg catacattgt 300
 gaacgaccac ccctgggatt ctggtgctga tggcgggtact tcggttcagg cggaggcatc 360
 cttaccaagg aatctgcttt tcaagtatgc caccaacagt gaagaggtta ttggagtgtat 420
 gagtaaagaa tacataccaa agggcacacg ttttggaacc ctaatagggtg aaatctacac 480
 caatgacaca gttcctaaga acgccaacag gaaatatatt tggaggatct attccagagg 540
 ggagcttcac cacttcattg acggctttta tgaagagaaa agcaactgga tgcgctatgt 600
 gaatccagca cactctcccc gggagcaaaa cctggctgcg tgtcagaacg ggatgaacat 660
 ctacttctac accattaagc ccattccctgc caaccaggaa cttcttgtgt ggtattgtcg 720
 ggactttgca gaaaggcttc actaccctta tcccggagag ctgacaatga tgaatctcac 780
 acaaacacag agcagtctaa agcaaccgag cactgagaaa aatgaactct gcccaaagaa 840
 tgtcccaaa agagagtaca gcgtgaaaga aatcctaaaa ttggactcca acccctccaa 900
 aggaaaggac ctctaccgtt ctaacatttc acccctcaca tcagaaaagg acctcgatga 960
 ctttagaaga cgtgggagcc ccgaaatgac cttctaccct cgggtcgttt accccatccg 1020
 ggcccctctg ccagaagact ttttgaaagc ttccctggcc tacgggatcg agagaccac 1080
 gtacatcact cgctcccca ttccatcttc caccactcca agcccctctg caagaagcag 1140
 ccccgaccaa agcctcaaga gctccagccc tcacagcagc cctgggaata cgggtgtccc 1200
 tgtgggcccc ggctctcaag agcaccggga ctctacgct tacttgaacg cgtcctacg 1260
 cacggaaggt ttgggtcctt accctggcta cgcacctctg cccacctcc cgccagcttt 1320
 catcccctcg tacaacgctc actaccctga gttcctcttg ccccccacg gcatgaattg 1380
 taatggcctg agcgtgtga gcagcatgaa ttgcatcaac aactttggcc tcttcccag 1440
 gctgtgccct gtctacagca atctcctcgg tgggggcagc ctgccccacc ccatgctcaa 1500
 cccacttct ctcccgagct cgctgccctc agatggagcc cggaggttg tccagccgga 1560
 gcatcccagg gaggtgcttg tcccggcgcc ccacagtgc ttctccttta ccggggccgc 1620
 cgccagcatg aaggacaagg cctgtagccc cacaagcggg tctcccacg cggaacagc 1680
 cgccagcgga gaacatgtgg tgcagcccaa agctacctca gcagcgatgg cagccccag 1740
 cagcgacgaa gccatgaatc tcattaaaaa caaaagaaac atgaccggct acaagaccct 1800
 tccttaccgg ctgaagaagc agaacggcaa gatcaagtac gaatgcaacg tttgcgcaa 1860
 gactttcggc cagctctcca atctgaaggc ccacctgaga gtgcacagtg gagaacggcc 1920
 tttcaaatgt cagacttgca acaagggtt tactcagctc gccacctgc agaaacacta 1980
 cctggtacac acgggagaaa agccacatga atgcccaggtc tgccacaaga gatttagcag 2040
 caccagcaat ctcaagacct acctgcgact ccattctgga gagaaacat accaatgcaa 2100
 ggtgtgccct gccaaagttc cccagtttgt gcacctgaaa ctgcacaagc gtctgcacac 2160
 ccgggagcgg cccacaagt gctcccagtg ccacaagaac tacatccatc tctgtagcct 2220
 caagggtcac ctgaaaggga actgcgctgc ggcccggcg cctgggctgc ccttggaaga 2280
 tctgaccgga atcaatgaag aaatcgagaa gtttgacatc agtgacaatg ctgaccggct 2340
 cgaggacgtg gaggatgaca tcagtgtgat ctctgtagtg gagaaggaaa ttctggccgt 2400
 ggtcagaaaa gagaaagaag aaactggcct gaaagtgtct ttgcaaagaa acatggggaa 2460
 tggactcctc tcctcagggt gcagccttta tgagtcatca gatctacccc tcatgaagtt 2520
 gcctcccagc aaccactac ctctggtacc tgtaaaggct aaacaagaaa cagttgaacc 2580
 aatggatcct taagattttc agaaaacact tattt 2615

<210> 120
 <211> 29
 <212> PRT
 <213> Homo sapiens

84

<400> 120

Leu Gln Asn Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly
1 5 10 15

Thr Cys Leu Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val
20 25

<210> 121

<211> 21

<212> ADN

<213> Homo sapiens

<400> 121

cttcaaaca caaccaggag g

21

<210> 122

<211> 20

<212> ADN

<213> Homo sapiens

<400> 122

ttggggaggt tggccgacga

20

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 99/01513

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/48 C12Q1/70 C07K14/15 A61K31/70

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C12Q C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 23755 A (BIO MERIEUX) 4 June 1998 (1998-06-04) Comparez nucléotides 1-1462 de SEQ ID NO:117 avec nucléotides 928-2390 de SEQ ID NO:1 de la présente demande; comparez SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de la présente demande ---	1,3-12, 14-36
X	Database GenBank. Séquence HSAC 000064 Clone humain BAC RG083M05 de 7q21-7q22, séquence complet. 17 novembre 1996. XP002118730 Comparez nucléotides 28000-38500 d' AC00064 avec SEQ ID NO:3 --- -/--	1-4,13

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- *Z* document member of the same patent family

Date of the actual completion of the international search

18 October 1999

Date of mailing of the international search report

11.11.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Cupido, M

INTERNATIONAL SEARCH REPORT

Application No

PCT/FR 99/01513

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>ALLIEL PM ET AL: "Séquences rétrovirales endogènes anlogues à celle du nouveau rétrovirus MSRV associé à la sclérose en plaques (première partie)"</p> <p>COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,</p> <p>vol. 321, no. 6, June 1998 (1998-06),</p> <p>pages 495-499, XP002101380</p> <p>MONTREUIL FR</p> <p>figures 2,3</p> <p>---</p>	1,3-12, 14-36
X	<p>Database GenBank Séquence AC X93499</p> <p>mRNA de H. sapiens pour la protéine rab7</p> <p>10 février 1997</p> <p>XP002119234</p> <p>& VITELLI R ET AL: "Molecular cloning and expression analysis of the human rab7 GTP-ase complementary deoxyribonucleic acid"</p> <p>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS,</p> <p>vol. 229, no. 3, 1996, pages 887-890,</p> <p>ORLANDO, FL US</p> <p>---</p>	1-4
X	<p>FR 2 737 500 A (BIO MERIEUX)</p> <p>7 February 1997 (1997-02-07)</p> <p>cited in the application</p> <p>the whole document</p> <p>---</p>	1,3-12, 14-36
P,X	<p>WO 99 02666 A (BIO MERIEUX)</p> <p>21 January 1999 (1999-01-21)</p> <p>Comparez SEQ ID NOs 130, 117, 114 et 120 avec SEQ ID NOs: 1-3 de cette demande et SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de cette demande</p> <p>---</p>	1,3-12, 14-36
P,X	<p>WO 99 02696 A (BIO MERIEUX ;BESEME FREDERIC (FR); BLOND JEAN LUC (FR); BOUTON OLI) 21 January 1999 (1999-01-21)</p> <p>Comparez SEQ ID NOs: 4, 5, 7, 9 et 11 avec SEQ ID NOs: 1-3 de cette demande.</p> <p>---</p>	1,3-12, 14-36
P,X	<p>ALLIEL PM ET AL: "Rétrovirus endogènes et sclérose en plaques. II. HERV-7q"</p> <p>COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,</p> <p>vol. 321, no. 10, October 1998 (1998-10),</p> <p>pages 857-863, XP002101381</p> <p>MONTREUIL FR</p> <p>the whole document</p> <p>---</p>	1,3-12, 14-36

-/--

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 99/01513

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 26972 A (GENETICS INSTITUTE, INC.) 3 June 1999 (1999-06-03) comparez SEQ ID N0:4 avec a.a. 131-668 de séquence 22 de cette demande ---	1-4
A	MITANI M ET AL: "Suppressive effect on polyclonal B-cell activation of a synthetic peptide homologous to a transmembrane component of oncogenic retrovirus" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 84, no. 1, January 1987 (1987-01), pages 237-240, XP002118729 WASHINGTON US cited in the application the whole document -----	24,25

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR 99/01513

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 8
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

See supplementary sheet INFORMATION FOLLOW-UP PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See supplementary sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

The International Searching Authority found several (groups of) inventions in the international application, namely:

1. Claims: 1, 9, 21-23, 26 (wholly), 28, 10-20 and 27-37 (partly)

Nucleic acid fragments derived from HERV-7q env, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

2. Claims: 2-7, 10-20, 27-37 (all partly)

Nucleic acid fragments derived from HERV-7q gag, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

3. Claims: 2-7, 10-20, 27-37 (all partly)

Human nucleic acid fragments similar to HERV-7q gag, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

4. Claims: 24, 25

Compositions comprising a CKS-type motif, inasmuch as said compositions do not contain a sequence as per the first invention.

Continuation of Box I.2

Claim No: 8

Claim 8 concerns a very wide variety of compounds. A support basis as defined in PCT Article 6 and a description as defined in PCT Article 5 can however be found for only a very limited number of the claimed compounds. In the present case, the claims are so lacking in support basis and the disclosure of the invention in the description is so limited that it is impossible to carry out any significant search concerning the whole claimed spectrum.

The applicant's attention is drawn to the fact that claims concerning inventions in respect of which no search report has been established need not be the subject of a preliminary examination report (PCT Rule 66.1 (e)). The applicant is warned that the guideline adopted by the EPO acting in its capacity as International Preliminary Examining Authority is not to proceed with a preliminary examination of a subject matter unless a search has been carried out thereon. This position will remain unchanged, notwithstanding that the claims have or have not been modified, either after receiving the search report, or during any procedure under Chapter II.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 99/01513

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9823755	A	04-06-1998	EP	0942987 A	22-09-1999
FR 2737500	A	07-02-1997	AU	6823296 A	05-03-1997
			BG	101355 A	30-12-1997
			BR	9606566 A	30-12-1997
			CA	2201282 A	20-02-1997
			CZ	9701357 A	17-06-1998
			EP	0789077 A	13-08-1997
			WO	9706260 A	20-02-1997
			HU	9900425 A	28-05-1999
			JP	11502416 T	02-03-1999
			NO	971493 A	03-06-1997
			NZ	316080 A	29-04-1999
			PL	319512 A	18-08-1997
			SK	56797 A	09-09-1998
WO 9902666	A	21-01-1999	FR	2765588 A	08-01-1999
			AU	8545098 A	08-02-1999
WO 9902696	A	21-01-1999	AU	8447098 A	08-02-1999
WO 9926972	A	03-06-1999	AU	1417899 A	15-06-1999

RAPPORT DE RECHERCHE INTERNATIONALE

Dem: Internationale No

PCT/FR 99/01513

A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 6 C12N15/48 C12Q1/70 C07K14/15 A61K31/70

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 C12N C12Q C07K A61K

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 98 23755 A (BIO MERIEUX) 4 juin 1998 (1998-06-04) Comparez nucléotides 1-1462 de SEQ ID NO:117 avec nucléotides 928-2390 de SEQ ID NO:1 de la présente demande; comparez SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de la présente demande	1,3-12, 14-36
X	Database GenBank. Séquence HSAC 000064 Clone humain BAC RG083M05 de 7q21-7q22, séquence complet. 17 novembre 1996. XP002118730 Comparez nucléotides 28000-38500 d' AC00064 avec SEQ ID NO:3	1-4,13

-/--



Voir la suite du cadre C pour la fin de la liste des documents



Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

18 octobre 1999

Date d'expédition du présent rapport de recherche internationale

11.11.99

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Cupido, M

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>ALLIEL PM ET AL: "Séquences rétrovirales endogènes anlogues à celle du nouveau rétrovirus MSRV associé à la sclérose en plaques (première partie)"</p> <p>COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,</p> <p>vol. 321, no. 6, juin 1998 (1998-06),</p> <p>pages 495-499, XP002101380</p> <p>MONTREUIL FR</p> <p>figures 2,3</p>	1,3-12, 14-36
X	<p>Database GenBank Séquence AC X93499</p> <p>mRNA de H. sapiens pour la protéine rab7</p> <p>10 février 1997</p> <p>XP002119234</p> <p>& VITELLI R ET AL: "Molecular cloning and expression analysis of the human rab7 GTP-ase complementary deoxyribonucleic acid"</p> <p>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS,</p> <p>vol. 229, no. 3, 1996, pages 887-890,</p> <p>ORLANDO, FL US</p>	1-4
X	<p>FR 2 737 500 A (BIO MERIEUX)</p> <p>7 février 1997 (1997-02-07)</p> <p>cité dans la demande</p> <p>le document en entier</p>	1,3-12, 14-36
P,X	<p>WO 99 02666 A (BIO MERIEUX)</p> <p>21 janvier 1999 (1999-01-21)</p> <p>Comparez SEQ ID NOs 130, 117, 114 et 120 avec SEQ ID NOs: 1-3 de cette demande et SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de cette demande</p>	1,3-12, 14-36
P,X	<p>WO 99 02696 A (BIO MERIEUX ; BESEME FREDERIC (FR); BLOND JEAN LUC (FR); BOUTON OLI) 21 janvier 1999 (1999-01-21)</p> <p>Comparez SEQ ID NOs: 4, 5, 7, 9 et 11 avec SEQ ID NOs: 1-3 de cette demande.</p>	1,3-12, 14-36
P,X	<p>ALLIEL PM ET AL: "Rétrovirus endogènes et sclérose en plaques. II. HERV-7q"</p> <p>COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,</p> <p>vol. 321, no. 10, octobre 1998 (1998-10),</p> <p>pages 857-863, XP002101381</p> <p>MONTREUIL FR</p> <p>le document en entier</p>	1,3-12, 14-36

-/--

RAPPORT DE RECHERCHE INTERNATIONALE

Dema Internationale No

PCT/FR 99/01513

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P,X	WO 99 26972 A (GENETICS INSTITUTE, INC.) 3 juin 1999 (1999-06-03) comparez SEQ ID NO:4 avec a.a. 131-668 de séquence 22 de cette demande -----	1-4
A	MITANI M ET AL: "Suppressive effect on polyclonal B-cell activation of a synthetic peptide homologous to a transmembrane component of oncogenic retrovirus" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 84, no. 1, janvier 1987 (1987-01), pages 237-240, XP002118729 WASHINGTON US cité dans la demande le document en entier -----	24,25

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR 99/01513

Cadre I Observation - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 1 de la première feuille)

Conformément à l'article-17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. ☐ Les revendications n°
se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:

2. ☒ Les revendications n° 8
se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:
voir feuille supplémentaire SUITE DES RENSEIGNEMENTS PCT/ISA/210

3. ☐ Les revendications n°
sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

Cadre II Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 2 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☐ Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.

2. ☒ Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.

3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n°

4. ☐ Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n°

Remarque quant à la réserve

- ☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.
- ☐ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

Suite du cadre I.2

Revendications nos.: 8

Le revendication 8 présente à trait à une très grande variété de composés. Un fondement au sens de L'Article 6 PCT et un exposé au sens de l'Article 5 PCT ne peut cependant être trouvé que pour un nombre très restreint de ces composés revendiqués. Dans le cas présent, les revendications manquent à un tel point de fondement et l'exposé de l'invention dans la description est si limité q'une recherche significative couvrant tout le spectre revendiqué est impossible.

L'attention du déposant est attirée sur le fait que les revendications, ou des parties de revendications, ayant trait aux inventions pour lesquelles aucun rapport de recherche n'a été établi ne peuvent faire obligatoirement l'objet d'un rapport préliminaire d'examen (Règle 66.1(e) PCT). Le déposant est averti que la ligne de conduite adoptée par l'OEB agissant en qualité d'administration chargée de l'examen préliminaire international est, normalement, de ne pas procéder à un examen préliminaire sur un sujet n'ayant pas fait l'objet d'une recherche. Cette attitude restera inchangée, indépendamment du fait que les revendications aient ou n'aient pas été modifiées, soit après la réception du rapport de recherche, soit pendant une quelconque procédure sous le Chapitre II.

SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: 1, 9, 21-23, 26 (complet), 28, 10-20 et 27-37 (partiellement)

Fragments d'acide nucléique dérivé du HERV-7q env, réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

2. revendications: 2-7, 10-20, 27-37(tous partiellement)

Fragments d'acide nucléique dérivé du HERV-7q gag, réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

3. revendications: 2-7, 10-20, 27-37(tous partiellement)

Fragments d'acide nucléique humaines similaires à HERV-7q (SEQ ID NOs: 4-21 et 61), réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

4. revendications: 24, 25

Compositions comprenant un motif de type CKS, dans la mesure où ces compositions ne contiennent pas une séquence selon la première invention.

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/FR 99/01513

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9823755	A	04-06-1998	EP 0942987 A	22-09-1999
FR 2737500	A	07-02-1997	AU 6823296 A	05-03-1997
			BG 101355 A	30-12-1997
			BR 9606566 A	30-12-1997
			CA 2201282 A	20-02-1997
			CZ 9701357 A	17-06-1998
			EP 0789077 A	13-08-1997
			WO 9706260 A	20-02-1997
			HU 9900425 A	28-05-1999
			JP 11502416 T	02-03-1999
			NO 971493 A	03-06-1997
			NZ 316080 A	29-04-1999
			PL 319512 A	18-08-1997
			SK 56797 A	09-09-1998
WO 9902666	A	21-01-1999	FR 2765588 A	08-01-1999
			AU 8545098 A	08-02-1999
WO 9902696	A	21-01-1999	AU 8447098 A	08-02-1999
WO 9926972	A	03-06-1999	AU 1417899 A	15-06-1999